AD_____

Award Number:
W81XWH-10-1-0090

TITLE:
Validation of Biomarkers Predictive of Recurrence Following
Prostatectomy

PRINCIPAL INVESTIGATOR:
Dr. Carlos Moreno

CONTRACTING ORGANIZATION:
Emory University
Atlanta, GA 30322

REPORT DATE:
April 2011

TYPE OF REPORT:
Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
               Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT:

     X Approved for public release; distribution unlimited

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE (DD-MM-YYYY)	2. REPORT TYPE	3. DATES COVERED (From - To)
14-04-2011	Annual	15 Mar 2010 – 14 Mar 2011

4. TITLE AND SUBTITLE
Validation of Biomarkers Predictive of Recurrence Following Prostatectomy

5a. CONTRACT NUMBER

5b. GRANT NUMBER
W81XWH-10-1-0090

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Dr. Carlos S. Moreno

cmoreno@emory.edu

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Emory University

Atlanta, GA 30322-1018

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command   Fort Detrick, Maryland 21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

Thus far we have made good progress on our goal to validate biomarkers of recurrence in prostate cancer. We have published a manuscript now in press at *The American Journal of Pathology* describing our set of biomarkers, and will be presenting these data orally at the American Society of Investigative Pathology annual meeting in Washington, DC on April 11, 2011. We have initiated collection of samples at the Atlanta VA Medical Center and begun tissue coring and RNA extraction. We have faced some administrative hurdles with getting the project started regarding IRB approvals and subcontract awards. However, now that all of these obstacles have been overcome, we expect rapid progress in the next year for data generation. In addition, the miRNA DASL platform has been discontinued by Illumina, and we are thus planning on transitioning the project to next generation sequencing methods, which are likely to be more accurate and comprehensive. Because of the delays in getting the project initiated, we are requesting a no cost one year extension to change the project from at two-year to a three-year project.

15. SUBJECT TERMS
Prostate Cancer, Biomarkers, Genes, MicroRNAs, Microarrays

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON USAMRMC
a. REPORT U	**b. ABSTRACT** U	**c. THIS PAGE** U	UU	46	**19b. TELEPHONE NUMBER** (include area code)

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39.18

Table of Contents

Introduction

An important challenge in prostate cancer research is to develop effective predictors of tumor recurrence following surgery in order to determine whether immediate adjuvant therapy is warranted. To identify biomarkers predictive of biochemical recurrence, we isolated the RNA from 70 formalin-fixed paraffin-embedded (FFPE) radical prostatectomy specimens with known long term outcome to perform DASL expression profiling with a custom-designed panel of 522 prostate cancer relevant genes that we designed. We identified a panel of ten protein-coding genes and two miRNA genes (RAD23B, FBP1, TNFRSF1A, CCNG2, NOTCH3, ETV1, BID, SIM2, ANXA1, miR-519d, and miR-647) that could be used to separate patients with and without biochemical recurrence (p < 0.001), as well as for the subset of 42 Gleason score 7 patients (p < 0.001). We performed an independent validation analysis on 40 samples and found that the biomarker panel was also significant at prediction of biochemical recurrence for all cases (p = 0.013) and for a subset of 19 Gleason score 7 cases (p = 0.010), both of which were adjusted for relevant clinical information including T-stage, PSA and Gleason score. Importantly, these biomarkers could significantly predict clinical recurrence for Gleason score 7 patients. A manuscript describing these biomarkers is now in press at the *American Journal of Pathology*[1].

In March 2010, we received funding for a two-year Prostate Cancer IDEA Award to validate and improve this set of biomarkers on an independent, large set of patient samples. Over the past year, we have dealt with a large number of administrative hurdles to obtain the necessary approvals to proceed with the research funded by this award. Now that all of these hurdles have been overcome, we are beginning to implement our research plan in earnest. Because of the delays in the beginning of the project, we are requesting a no-cost extension of the project from a 2-year to a 3-year project.

Body

Custom Prostate DASL profiling

Funding for this IDEA Award was based on our DASL expression profiling data using our custom-designed prostate cancer panel and the Illumina DASL microRNA (miRNA) panel on 70 prostatectomy patient samples to identify biomarkers predictive of recurrence. In addition, an independent validation profiling experiment was performed on 40 additional samples. MiRNA probes were filtered to retain only those that were present on the miRNA microarrays used for both the training and validation sets, reducing the total number of probes examined to 403 miRNA probes. The training set included 29 cases with observed biochemical PSA recurrence (median time to recurrence = 19 months), and 41 cases censored, i.e., without observed recurrence during follow-up (median follow-up time = 83.0 months). A summary of the clinical characteristics of the training and validation sets of samples is provided in **Table 1**.

Integrated DASL biomarker analysis

After fitting a univariate Cox proportional hazard (PH) model for each individual probe using the training data, a set of 27 important probes were preselected based on an FDR threshold of 0.30. Next, to identify the optimal prediction score based on the preselected probes, we fit a lasso Cox proportional hazard (PH) model[2, 3] first using the set of 25 preselected mRNA probes only, resulting in a panel of nine protein-coding genes shown in **Table 2** (*RAD23B, FBP1, TNFRSF1A, NOTCH3, ETV1, BID, SIM2, ANXA1*, and *BCL2*). A final prediction model was then built to include the predictive score based on this panel of nine mRNA biomarkers as well as the relevant clinical biomarkers including T-stage, PSA and Gleason score, which could be used to predict recurrence following radical prostatectomy. Kaplan-Meier analysis (**Figure 1A**) demonstrated that these probes could significantly discriminate patients at higher and lower risk of recurrence by the log rank test ($p < 0.001$). We next applied the final predictive model developed on the training set to the validation set, a separate, independent DASL profiling experiment performed on a different day. Kaplan-Meier analysis (**Figure 1B**) on this validation set determined that the model could discriminate patients at higher and lower risk of recurrence ($p = 0.010$).

Subsequently, we repeated the above training procedure using the complete set of 27 preselected mRNA and miRNA probes, and we identified an optimal panel of ten mRNAs and two microRNAs (**Table 3**) and built a final prediction model for prostate cancer biochemical recurrence, which again included relevant clinical biomarkers. Kaplan-Meier analysis and the log-rank test determined that this panel could also significantly discriminate patients at higher and lower risk of recurrence both in the training set ($p < 0.001$, **Figure 1C**) and in the validation set ($p = 0.013$, **Figure 1D**).

Prediction of Cases with a Gleason Score 7

Prediction of recurrence for patients with a Gleason score 7 is particularly difficult. In order to address this issue, we applied the biomarker panels to the subset of cases in the training and validation sets that had a Gleason score 7. The prediction model based on the nine-mRNA panel was significant at discriminating biochemical recurrence in Gleason score 7 cases in both the training set ($p < 0.001$, Figure 2A) and the validation set ($p = 0.027$, Figure 2B). For the prediciton model based on the combined panel of ten mRNAs and two miRNAs in Table 3, the predictive value was again significant for both the training set ($p = < 0.001$, Figure 2C) and the validation set ($p = 0.010$, Figure 2D). A summary of the p-values for predicting biochemical recurrence is given in **Table 4**. In all cases, the prediction models that use one of the two gene biomarker panels plus clinical information outperforms the prediction model using only clinical information.

Analysis of clinical recurrence

Although most patients who have clinical recurrence following prostatectomy also have

biochemical recurrence, there is a significant population of patients with biochemical recurrence who do not have clinically significant recurrences observed during their follow-ups. To evaluate our biomarker panel of biochemical recurrence for predicting the clinical recurrence, we tested the prediction model based on the combined mRNA/miRNA panel in the same training and validation samples using their clinical recurrence outcome data. Unfortunately, clinical recurrence data was lacking on some of the samples, and the total number of samples used in the training set was reduced. In the training data, the combined mRNA/miRNA panel was highly significant for predicting clinical recurrence in all patients (p=0.002) as well as in the subset of patients with a Gleason score 7 (p=0.004); in the validation data, it was also significant for predicting recurrence in patients with a Gleason score 7 (p=0.023) and trended towards significance in all patients (p=0.078). A summary of the p-values for predicting clinical recurrence is given in **Table 5**. In all cases, the prediction model that uses the combined mRNA and miRNA panel plus the clinical information, again, outperforms the prediction model that uses only the clinical information.

We also performed an analysis to construct a predictive set of biomarkers based on the clinical recurrence data instead of biochemical recurrence. Only three probes passed the initial preselection step for the univariate Cox PH modeling, all corresponding to the *ETV1* gene, which is likely due to the considerably fewer number of clinical recurrences in the training set as well as the smaller total sample size. Furthermore, the prediction model built on this set of gene biomarkers did not perform as well as the models built on biochemical recurrence (data not shown).

American Journal of Pathology Manuscript

A manuscript describing these results was submitted to the *American Journal of Pathology* in September, 2010, a revision was submitted in December, 2010, and the paper was accepted for publication in March, 2011. This manuscript is attached as Appendix 1.

TABLES

	Training Set (Total)	Training Set (No BCR)	Training Set (BCR)	Validation Set (Total)	Validation Set (No BCR)	Validation Set (BCR)
Number Cases	70	41	29	40	27	13
Clinical Recurrence	8	0	8	11	0	11
No Clinical Recurrence	57	41	16	29	27	2
Median Time F/U (months)	84	83	81	74	75	73
Median Time to BCR (months)	19	N/A	19	14	N/A	14
Median Time no BCR (months)	48	83	19	34.5	56	14
Gleason Score (Avg +/- SD)	6.9 +/- 0.6	6.7 +/- 0.6	7.0 +/- 0.6	7.0 +/- 0.8	6.8 +/- 0.7	7.4 +/- 1
PSA (Avg +/- SD)	9.2 +/- 5.4	8.7 +/- 6.4	9.9 +/- 3.8	12.7 +/- 8.4	12.4 +/- 9.9	13.1 +/- 5.3
Age (Avg +/- SD)	61.9 +/- 7.7	61.2 +/- 7.7	62.9 +/- 7.8	63.6 +/- 8.4	63.5 +/- 8.3	64 +/- 8.9

Table 1: A summary of the clinical characteristics of the training and validation sets of patient samples. (BCR = Biochemical Recurrence, F/U = follow up, PSA = prostate specific antigen, SD = standard deviation).

Symbol	Description	Coefficient	References
RAD23B	RAD23 homolog B	0.152	[4, 5]
FBP1	Fructose-1,6-bisphosphatase 1	0.310	[6-8]
TNFRSF1A	Tumor Necrosis Factor Receptor Superfamily, Member 1A	-0.560	[9, 10]
NOTCH3	Notch homolog 3	0.426	[11, 12]
ETV1	Ets Variant Gene 1 (ETV1)	0.157	[13, 14]
BID	BH3 Interacting Domain Death Agonist (BID)	0.248	[15, 16]
SIM2	Single-Minded Homolog 2	0.043	[17-20]
ANXA1	Annexin A1	-0.185	[21-24]
BCL2	B-cell CLL/lymphoma 2	0.028	[25, 26]

Table 2: Nine-gene predictor of prostate cancer recurrence following surgery. Coefficient is derived from the lasso Cox proportion hazards model and was used for computing the predictive score. Positive coefficients indicate a positive association with recurrence, and negative coefficients a negative association with recurrence.

Symbol	Description	Coefficient	References
RAD23B	RAD23 homolog B	0.070	[4, 5]
FBP1	Fructose-1,6-bisphosphatase 1	0.251	[6-8]
TNFRSF1A	Tumor necrosis factor receptor superfamily, member 1A	-0.588	[9, 10]
CCNG2	Cyclin G2	0.008	[27-29]
hsa-miR-647	hsa-miR-647	-0.318	
LETMD1	LETM1 domain containing 1	0.063	[30-33]
NOTCH3	Notch homolog 3	0.367	[11, 12]
ETV1	ETS variant gene 1 (ETV1)	0.179	[13, 14]
hsa-miR-519d	hsa-miR-519d	0.551	[34]
BID	BH3 interacting domain death agonist (BID)	0.128	[15, 16]
SIM2	Single-minded homolog 2	0.124	[17-20]
ANXA1	Annexin A1	-0.143	[21-24]

Table 3: Twelve-gene predictor of prostate cancer recurrence following surgery using ten mRNAs and two microRNAs. Coefficient is derived from the lasso Cox proportion hazards model and was used for computing the predictive score. Positive coefficients indicate a positive association with recurrence, and negative coefficients a negative association with recurrence.

Training Set	mRNA panel	Combined mRNA/miRNA panel	Clinical Information Only
All Cases (n=61)	**<0.001**	**<0.001**	0.096
Gleason score 7 (n=42)	**<0.001**	**<0.001**	0.641
Validation Set	mRNA panel	Combined mRNA/miRNA panel	Clinical Information Only
All Cases (n=35)	0.010	**0.013**	**0.020**
Gleason score 7 (n=19)	0.027	**0.010**	**0.028**

Table 4: Summary of p-values (Logrank test) of prediction of biochemical recurrence on training and validation sets for the entire dataset and the subset of Gleason score 7 cases using two biomarker panels, all of which are adjusted forT-stage, PSA, and Gleason score, or using clinical information only. Significant p-values are indicated in bold.

Training Set	Combined mRNA/miRNA panel	Clinical Information Only
All Cases (n=56)	**0.002**	0.262
Gleason score 7 (n=37)	**0.004**	0.136
Validation Set	Combined mRNA/miRNA panel	Clinical Information Only
All Cases (n=35)	0.078	0.193
Gleason score 7 (n=19)	**0.023**	0.080

Table 5: Summary of p-values (Logrank test) of prediction of clinical recurrence on training and validation sets for the entire dataset and the subset of Gleason score 7 cases using the combined and mRNA/miRNA panel, all of which are adjusted for T-stage, PSA, and Gleason score, or using clinical information only. Significant p-values are indicated in bold.
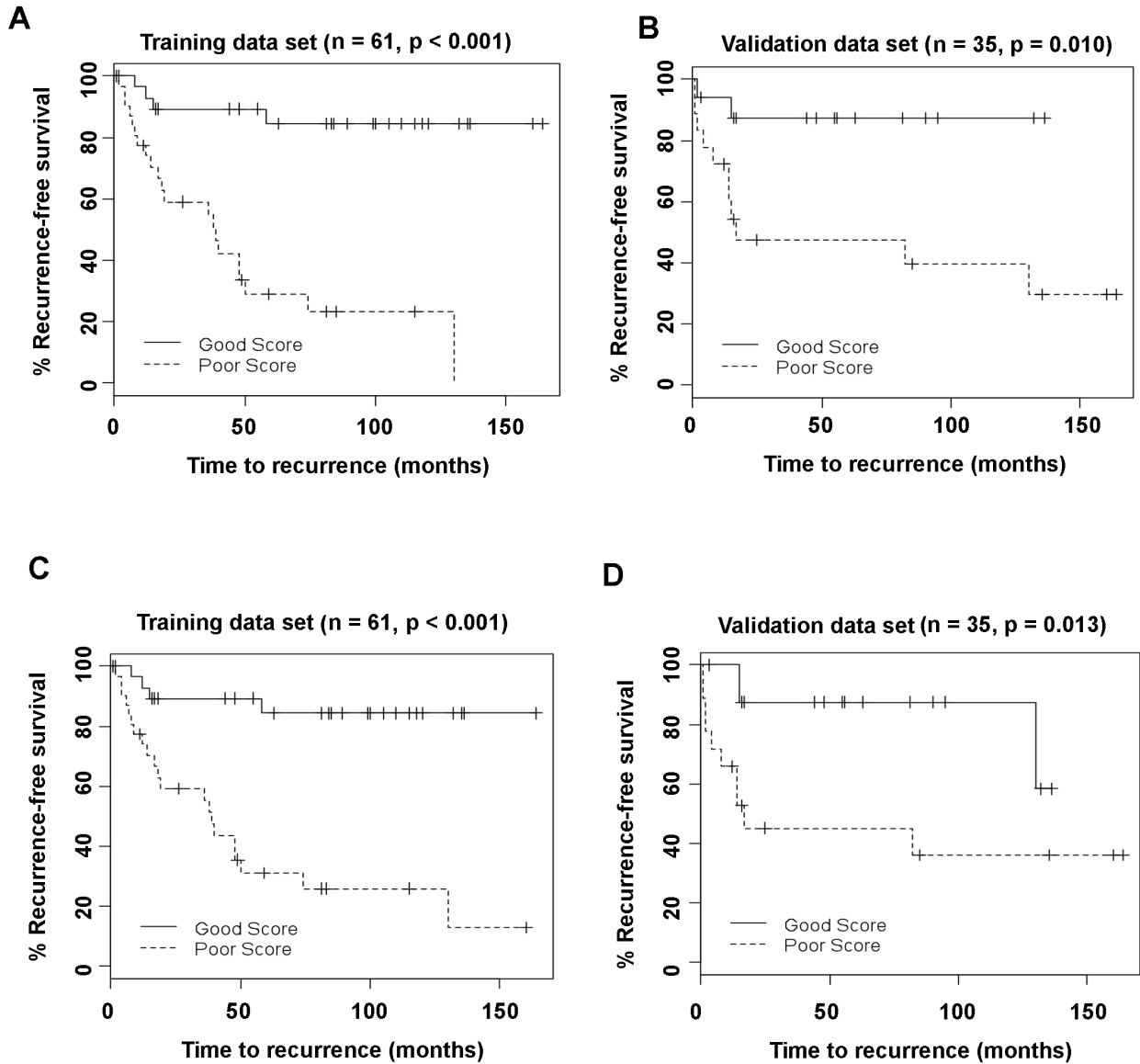
FIGURES

Figure 1: **Prediction of biochemical recurrence in all prostate cancer patients using two biomarker panels, adjusted for clinical information.** **(A)** Kaplan-Meier analysis of the training set patients that were separated based on the mRNA panel described in Table 2. **(B)** Kaplan-Meier analysis on the validation cases using the mRNA panel. **(C)** Kaplan-Meier analysis of the training set using the combined mRNA and miRNA panel described in Table 3. **(D)** Kaplan-Meier analysis of the validation set using the combined mRNA and miRNA panel.
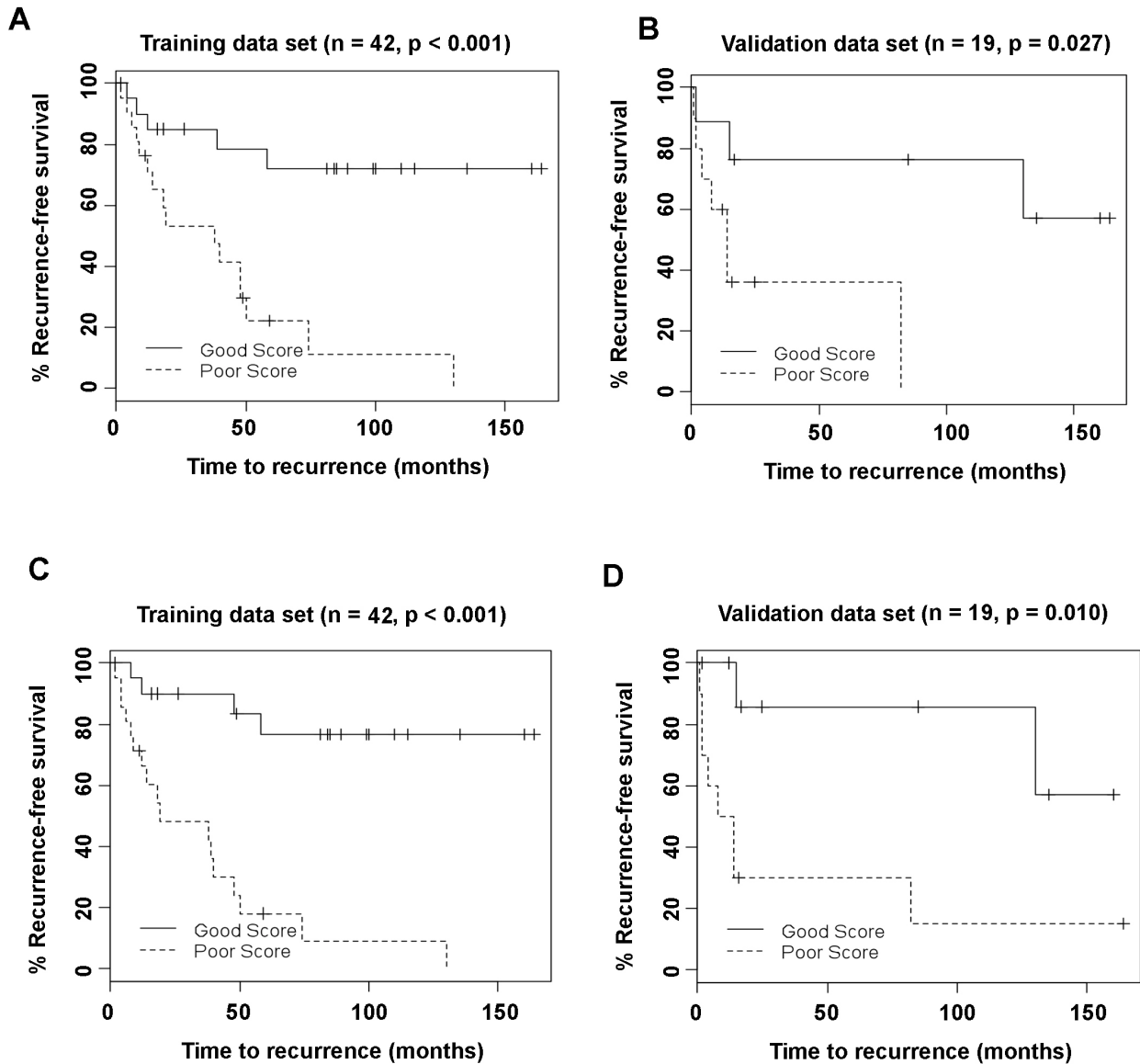
Figure 2: Prediction of biochemical recurrence in prostate cancer patients with a Gleason score 7 using two biomarker panels, adjusted for clinical information. (**A**) Kaplan-Meier analysis of the training set of Gleason score 7 cases using the mRNA panel described in Table 2. (**B**) Kaplan-Meier analysis of the Gleason score 7 cases in the validation set using the mRNA panel. (**C**) Kaplan-Meier analysis of the Gleason score 7 cases in the training set using the combined mRNA and miRNA panel described in Table 3. (**D**) Kaplan-Meier analysis of the Gleason score 7 cases in the validation set using the combined mRNA and miRNA panel.

Administrative hurdles

When this project was initiated, we anticipated that the IRB protocols already in place would be sufficient for conducting the described research project. However, review by the DOD determined that the IRB protocols were too general and not sufficiently specific. Consequently, we submitted new, specific IRB protocols at Emory University, the Emory/Atlanta VA Medical Center, and at Sunnybrook Research Centre at the University of Toronto. While we obtained IRB approval letters fairly quickly at Emory and the VA, approval from Sunnybrook took much longer. The difficulty hinged on award of the subcontract, since the Sunnybrook IRB would not issue approval without funding, and Emory would not issue the subcontract without the IRB approval. We were stuck in a Catch-22 for several months before we finally obtained Sunnybrook IRB approval in November, 2010 and permission to begin work at Sunnybrook was obtained in December, 2010. Unfortunately, it took another two months before all of the legal issues could be resolved, including the fact that there is no HIPAA law in Canada, so that the subcontract could be awarded to Sunnybrook by Emory University. Now that all of the administrative hurdles have been overcome, we expect to begin receiving samples from Toronto in the very near future.

Progress at Emory University

In the meantime, once approval was obtained from the DOD on August 13, 2010 to commence work on the project at Emory, we began immediately to work on identifying samples that we could use in our validation study. We identified 150 cases at the VA hospital between 1990-2000 that could potentially be used for this project. We were able to locate slides and formalin-fixed paraffin embedded (FFPE) blocks for 100 of those cases, and Dr. Oyesiku identified regions of cancer and benign tissue in slides for each of them. These samples were then submitted for processing to obtain 1 mm tissue cores. We anticipate that we will have RNA ready for WG-DASL analysis in the next few weeks. We are planning to examine additional cases at Emory and the VA between 2000-2003 to identify samples for use in Aim 2.

Platform issues

In our initial proposal, we planned to use the Illumina miRNA DASL platform for analysis of miRNA biomarker expression levels. Since submission of our application for initial review, Illumina has discontinued this platform. Consequently, we are evaluated several different options for comprehensive analysis of miRNA levels. These options include TaqMan Low Density Arrays, Affymetrix miRNA arrays, High-Throughput Genomics Quantitative Nuclease Protection Assays, Illumina HiSeq sequencing, Ion Torrent sequencing, and Nanostring sequencing. There are strengths and weaknesses to each of these options that we are currently scrutinizing. However, we are likely to use a 48-plex multiplexing method of next generation sequencing on the Illumina HiSeq platform. This should give us deep coverage and high data quality at an acceptable cost. We plan on performing pilot studies to determine whether we can obtain acceptable data using FFPE-derived RNA shortly.

Key Research Accomplishments

Our key research accomplishments are summarized below:

- Identified a set of 10 mRNAs and 2 miRNAs predictive of recurrence following prostatectomy.

- Published a manuscript describing these biomarkers of recurrence.

- Obtained all necessary IRB and DOD approvals to commence work.

- Identified 100 prostate cancer cases at the Atlanta/VA for the validation study.

- Marked tumor and benign areas to enable coring of FFPE blocks.

- Initiated RNA extraction of the first 100 cases for WG-DASL analysis.

Reportable Outcomes

Grant Award Received	March 25, 2010
New IRB Protocol Requested from DOD	May 20, 2010
New IRB Protocol Approved by Emory University	June 20, 2010
Approval received from VA Research Committee to use VA samples	August 10, 2010
Permission to begin research at Emory obtained from DOD	August 13, 2010
Manuscript describing initial biomarkers submitted for publication	September 24, 2010
Initial set of 100 samples identified at the Emory/VA	October 12, 2010
Abstract submitted to American Society for Investigative Pathology (ASIP) 2011 meeting in Washington, DC	November 5, 2010
New IRB Protocol Approved by U. Toronto/Sunnybrook	November 15, 2010
Permission to begin research at U.Toronto/Sunnybrook obtained from DOD	December 21, 2010
Revised manuscript describing biomarkers submitted for publication	December 31, 2010
Abstract for ASIP 2011 meeting selected for oral presentation	February 7, 2011
Tissue blocks pulled and slides marked for 100 samples at the VA	February 18, 2010
Subaward Contract agreed between Emory and U. Toronto/Sunnybrook	February 28, 2011
Manuscript describing initial biomarkers accepted for publication in *AJP*[1]	March 3, 2011
Tissue coring from 100 VA FFPE tissue blocks initiated	March 15, 2011

Conclusion

Thus far we have made good progress on our goal to validate biomarkers of recurrence in prostate cancer. We have published a manuscript now in press at *The American Journal of Pathology* describing our set of biomarkers[1], and will be presenting these data orally at the American Society of Investigative Pathology annual meeting in Washington, DC on April 11, 2011. We have initiated collection of samples at the Atlanta VA Medical Center and begun tissue coring and RNA extraction. We have faced some administrative hurdles with getting the project started regarding IRB approvals and subcontract awards. However, now that all of these obstacles have been overcome, we expect rapid progress in the next year for data generation. In addition, the miRNA DASL platform has been discontinued by Illumina, and we are thus planning on transitioning the project to next generation sequencing methods, which are likely to be more accurate and comprehensive. Because of the delays in getting the project initiated, we are requesting a no cost one year extension to change the project from at two-year to a three-year project.

References

1. Long Q, Johnson BA, Osunkoya AO, Lai Y-H, Zhou W, Abramovitz M, Xia M, Bouzyk M, Nam RK, Sugar L, Stanimirovic A, Leyland-Jones B, Seth AK, Petros JA, Moreno CS: Protein-coding and MicroRNA Biomarkers of Recurrence of Prostate Cancer Following Radical Prostatectomy, Am J Pathol 2011, in press
2. Tibshirani R: The lasso method for variable selection in the Cox model, Stat Med 1997, 16:385-395
3. Goeman JJ: L1 penalized estimation in the Cox proportional hazards model, Biom J 2010, 52:70-84
4. Thoma BS, Vasquez KM: Critical DNA damage recognition functions of XPC-hHR23B and XPA-RPA in nucleotide excision repair, Mol Carcinog 2003, 38:1-13
5. Brignone C, Bradley KE, Kisselev AF, Grossman SR: A post-ubiquitination role for MDM2 and hHR23A in the p53 degradation pathway, Oncogene 2004, 23:4121-4129
6. Yanez AJ, Nualart F, Droppelmann C, Bertinat R, Brito M, Concha, II, Slebe JC: Broad expression of fructose-1,6-bisphosphatase and phosphoenolpyruvate carboxykinase provide evidence for gluconeogenesis in human tissues other than liver and kidney, J Cell Physiol 2003, 197:189-197
7. Lapointe J, Li C, Higgins JP, Van De Rijn M, Bair E, Montgomery K, Ferrari M, Egevad L, Rayford W, Bergerheim U, Ekman P, DeMarzo AM, Tibshirani R, Botstein D, Brown PO, Brooks JD, Pollack JR: Gene expression profiling identifies clinically relevant subtypes of prostate cancer, Proc Natl Acad Sci U S A 2004, 101:811-816
8. Liu P, Ramachandran S, Ali Seyed M, Scharer CD, Laycock N, Dalton WB, Williams H, Karanam S, Datta MW, Jaye DL, Moreno CS: Sex-determining region Y box 4 is a transforming oncogene in human prostate cancer cells, Cancer Res 2006, 66:4011-4019
9. Thorburn A: Death receptor-induced cell killing, Cell Signal 2004, 16:139-144
10. Hsu H, Xiong J, Goeddel DV: The TNF receptor 1-associated protein TRADD signals cell death and NF-kappa B activation, Cell 1995, 81:495-504
11. Wang Z, Li Y, Ahmad A, Banerjee S, Azmi AS, Kong D, Wojewoda C, Miele L, Sarkar FH: Down-regulation of Notch-1 is associated with Akt and FoxM1 in inducing cell growth inhibition and apoptosis in prostate cancer cells, J Cell Biochem 2010,
12. Wang Z, Li Y, Banerjee S, Kong D, Ahmad A, Nogueira V, Hay N, Sarkar FH: Down-regulation of Notch-1 and Jagged-1 inhibits prostate cancer cell growth, migration and invasion, and induces apoptosis via inactivation of Akt, mTOR, and NF-kappaB signaling pathways, J Cell Biochem 2010, 109:726-736
13. Tomlins SA, Rhodes DR, Perner S, Dhanasekaran SM, Mehra R, Sun XW, Varambally S, Cao X, Tchinda J, Kuefer R, Lee C, Montie JE, Shah RB, Pienta KJ, Rubin MA, Chinnaiyan AM: Recurrent fusion of TMPRSS2 and ETS transcription factor genes in prostate cancer, Science 2005, 310:644-648
14. Tomlins SA, Laxman B, Dhanasekaran SM, Helgeson BE, Cao X, Morris DS, Menon A, Jing X, Cao Q, Han B, Yu J, Wang L, Montie JE, Rubin MA, Pienta KJ, Roulston D, Shah RB, Varambally S, Mehra R, Chinnaiyan AM: Distinct classes of chromosomal rearrangements create oncogenic ETS gene fusions in prostate cancer, Nature 2007, 448:595-599
15. Kulik G, Carson JP, Vomastek T, Overman K, Gooch BD, Srinivasula S, Alnemri E, Nunez G, Weber MJ: Tumor necrosis factor alpha induces BID cleavage and bypasses antiapoptotic signals in prostate cancer LNCaP cells, Cancer Res 2001, 61:2713-2719
16. Luo X, Budihardjo I, Zou H, Slaughter C, Wang X: Bid, a Bcl2 interacting protein, mediates cytochrome c release from mitochondria in response to activation of cell surface death receptors, Cell 1998, 94:481-490
17. Arredouani MS, Lu B, Bhasin M, Eljanne M, Yue W, Mosquera JM, Bubley GJ, Li V, Rubin MA, Libermann TA, Sanda MG: Identification of the transcription factor single-minded homologue 2 as a potential biomarker and immunotherapy target in prostate cancer, Clin Cancer Res 2009, 15:5794-5802

18. Halvorsen OJ, Rostad K, Oyan AM, Puntervoll H, Bo TH, Stordrange L, Olsen S, Haukaas SA, Hood L, Jonassen I, Kalland KH, Akslen LA: Increased expression of SIM2-s protein is a novel marker of aggressive prostate cancer, Clin Cancer Res 2007, 13:892-897

19. DeYoung MP, Tress M, Narayanan R: Identification of Down's syndrome critical locus gene SIM2-s as a drug therapy target for solid tumors, Proc Natl Acad Sci U S A 2003, 100:4760-4765

20. Deyoung MP, Scheurle D, Damania H, Zylberberg C, Narayanan R: Down's syndrome-associated single minded gene as a novel tumor marker, Anticancer Res 2002, 22:3149-3157

21. Maschler S, Gebeshuber CA, Wiedemann EM, Alacakaptan M, Schreiber M, Custic I, Beug H: Annexin A1 attenuates EMT and metastatic potential in breast cancer, EMBO Mol Med 2010,

22. Inokuchi J, Lau A, Tyson DR, Ornstein DK: Loss of annexin A1 disrupts normal prostate glandular structure by inducing autocrine IL-6 signaling, Carcinogenesis 2009, 30:1082-1088

23. Patton KT, Chen HM, Joseph L, Yang XJ: Decreased annexin I expression in prostatic adenocarcinoma and in high-grade prostatic intraepithelial neoplasia, Histopathology 2005, 47:597-601

24. Paweletz CP, Ornstein DK, Roth MJ, Bichsel VE, Gillespie JW, Calvert VS, Vocke CD, Hewitt SM, Duray PH, Herring J, Wang QH, Hu N, Linehan WM, Taylor PR, Liotta LA, Emmert-Buck MR, Petricoin EF, 3rd: Loss of annexin 1 correlates with early onset of tumorigenesis in esophageal and prostate carcinoma, Cancer Res 2000, 60:6293-6297

25. Alnemri ES, Robertson NM, Fernandes TF, Croce CM, Litwack G: Overexpressed full-length human BCL2 extends the survival of baculovirus-infected Sf9 insect cells, Proc Natl Acad Sci U S A 1992, 89:7295-7299

26. Wagner AJ, Small MB, Hay N: Myc-mediated apoptosis is blocked by ectopic expression of Bcl-2, Mol Cell Biol 1993, 13:2432-2440

27. Arachchige Don AS, Dallapiazza RF, Bennin DA, Brake T, Cowan CE, Horne MC: Cyclin G2 is a centrosome-associated nucleocytoplasmic shuttling protein that influences microtubule stability and induces a p53-dependent cell cycle arrest, Exp Cell Res 2006, 312:4181-4204

28. Martinez-Gac L, Marques M, Garcia Z, Campanero MR, Carrera AC: Control of cyclin G2 mRNA expression by forkhead transcription factors: novel mechanism for cell cycle control by phosphoinositide 3-kinase and forkhead, Mol Cell Biol 2004, 24:2181-2189

29. Bates S, Rowan S, Vousden KH: Characterisation of human cyclin G1 and G2: DNA damage inducible genes, Oncogene 1996, 13:1103-1109

30. Cho GW, Kim MH, Kim SH, Ha SA, Kim HK, Kim S, Kim JW: TCF/beta-catenin plays an important role in HCCR-1 oncogene expression, BMC Mol Biol 2009, 10:42

31. Cho GW, Shin SM, Namkoong H, Kim HK, Ha SA, Hur SY, Kim TE, Chai YG, Kim JW: The phosphatidylinositol 3-kinase/Akt pathway regulates the HCCR-1 oncogene expression, Gene 2006, 384:18-26

32. Jung SS, Park HS, Lee IJ, Namkoong H, Shin SM, Cho GW, Ha SA, Park YG, Lee YS, Ko J, Kim JW: The HCCR oncoprotein as a biomarker for human breast cancer, Clin Cancer Res 2005, 11:7700-7708

33. Ko J, Lee YH, Hwang SY, Lee YS, Shin SM, Hwang JH, Kim J, Kim YW, Jang SW, Ryoo ZY, Kim IK, Namkoong SE, Kim JW: Identification and differential expression of novel human cervical cancer oncogene HCCR-2 in human cancers and its involvement in p53 stabilization, Oncogene 2003, 22:4679-4689

34. Martinelli R, Nardelli C, Pilone V, Buonomo T, Liguori R, Castano I, Buono P, Masone S, Persico G, Forestieri P, Pastore L, Sacchetti L: miR-519d Overexpression Is Associated With Human Obesity, Obesity (Silver Spring) 2010,

Appendices

See attached manuscript by Long et al[1].

Protein-coding and MicroRNA Biomarkers of Recurrence of Prostate Cancer Following Radical Prostatectomy

Qi Long[1,4], Brent A. Johnson[1,], Adeboye O. Osunkoya[3,4,5,6], Yu-Heng Lai[2,3], Wei Zhou[2,3,4,7,9], Mark Abramovitz[8,11], Mingjing Xia[4], Mark B. Bouzyk[4,9], Robert K. Nam[10], Linda Sugar[10], Aleksandra Stanimirovic[10], Brian R. Leyland-Jones[4,7], Arun K. Seth[10], John A. Petros[3,4,5,6,7], and Carlos S. Moreno[2,3,4]

[1]*Department of Biostatistics and Bioinformatics, Emory University*
[2]*Program in Genetics and Molecular Biology, Emory University,*
[3]*Department of Pathology and Laboratory Medicine, Emory University,*
[4]*Winship Cancer Institute,*
[5]*Urology Department, Emory University,*
[6]*Atlanta VA Medical Center, Decatur, Georgia.*
[7]*Department of Hematology and Medical Oncology, Emory University,*
[8]*VM Institute of Research, Montreal, Quebec,*
[9]*Department of Human Genetics, Emory University,*
[10]*Department of Laboratory Medicine and Pathobiology, University of Toronto, and Department of Anatomic Pathology, Sunnybrook Health Sciences Centre*
[11] *Current address: Departments of Pathology & Oncology, Jewish General Hospital, 3755 Cote Ste Catherine, Montreal, QC, H3T 1E2; Canada*

Corresponding Author:
Carlos S. Moreno
Associate Professor
Department of Pathology & Laboratory Medicine
Winship Cancer Institute
Emory University

Whitehead Research Building, Rm. 105J
615 Michael St.
Atlanta GA, 30322

(404) 712-2809 (Ph)
(404) 727-8538 (Fax)
cmoreno@emory.edu

Key Words: Prostate Cancer, Recurrence, Biomarkers

Running Title: Biomarkers of Recurrence in Prostate Cancer

ABSTRACT

An important challenge in prostate cancer research is to develop effective predictors of tumor recurrence following surgery in order to determine whether immediate adjuvant therapy is warranted. To identify biomarkers predictive of biochemical recurrence, we isolated the RNA from 70 formalin-fixed paraffin-embedded (FFPE) radical prostatectomy specimens with known long term outcome to perform DASL expression profiling with a custom-designed panel of 522 prostate cancer relevant genes that we designed. We identified a panel of ten protein-coding genes and two miRNA genes (*RAD23B, FBP1, TNFRSF1A, CCNG2, NOTCH3, ETV1, BID, SIM2, ANXA1, miR-519d,* and *miR-647*) that could be used to separate patients with and without biochemical recurrence (p < 0.001), as well as for the subset of 42 Gleason score 7 patients (p < 0.001). We performed an independent validation analysis on 40 samples and found that the biomarker panel was also significant at prediction of recurrence for all cases (p = 0.013) and for a subset of 19 Gleason score 7 cases (p = 0.010), both of which were adjusted for relevant clinical information including T-stage, PSA and Gleason score. Importantly, these biomarkers could significantly predict clinical recurrence for Gleason score 7 patients. These biomarkers may increase the accuracy of prognostication following radical prostatectomy using formalin-fixed specimens.

INTRODUCTION

Prostate cancer remains the most common non-cutaneous cancer diagnosed for U.S. males, and ranks second among tumor site-specific mortality, with estimates for 2009 at over 192,000 new cases and 27,000 deaths[1]. The majority of patients with prostate cancer are clinically asymptomatic with early-stage, organ-confined disease, and in fact, more than 50% of men who reach the age of 80 develop clinically insignificant prostate cancer. However, a subpopulation of prostate cancer patients progress to highly aggressive, androgen-independent metastatic disease, which is inevitably fatal. One of the important challenges in current prostate cancer research is to develop effective methods to determine whether a patient is likely to progress to aggressive, metastatic disease in order to aid clinicians in deciding on the appropriate course of treatment. Biomarker assays that could predict progression and metastasis for prostate cancer patients would be of great utility in aiding clinical management of this large patient population. An important challenge in prostate cancer research is to develop effective predictors of tumor recurrence following surgery in order to determine whether immediate adjuvant therapy is warranted. Thus, biomarkers that could predict the likelihood of success for surgical therapies would be of great clinical significance.

In the past few years, enormous progress has been made in developing technologies to exploit formalin-fixed paraffin-embedded (FFPE) tumor tissue samples for gene expression analysis. The DASL (cDNA-mediated Annealing, Selection, extension and Ligation) assay is a unique expression profiling platform that is based upon massively multiplexed RT-PCR applied in a microarray format, that allows for the determination of expression of RNA isolated from 96 FFPE tumor tissue samples in a high throughput format [2, 3].

Here, we have identified biomarkers predictive of recurrence by expression profiling archived FFPE tumor samples using both a custom panel of prostate cancer associated mRNA genes and a panel

of microRNA genes. These biomarkers were developed on a training set of 70 patients (29 with biochemical recurrence and 41 controls) and validated on an independent set of 40 samples (13 with biochemical recurrence and 27 controls) and were able to significantly discriminate between patients with and without biochemical recurrence following radical prostatectomy. Moreover, these biomarkers were able to discriminate biochemical recurrence in patients with Gleason score 7, for whom outcome is particularly difficult to predict.

MATERIALS AND METHODS

Patient Samples

In the initial training set, 70 cases were used (29 with biochemical recurrence and 41 controls), 45 patients from Sunnybrook Health Science Center (Toronto, ON), and 25 patients from Emory University. The 45 cases of paraffin-embedded tissue samples from Toronto were drawn from men who underwent radical prostatectomy as the sole treatment for clinically localized prostate cancer (PCa) between 1998 and 2006. The clinical data includes multiple clinicopathologic variables such as prostate specific antigen (PSA) levels, histologic grade (Gleason score), tumor stage (pathologic stage category for example; organ confined, pT2; or with extra-prostatic extension, pT3a; or with seminal vesicle invasion, pT3b), and biochemical recurrence rates. For the cases from Emory University, both the training set (25 cases) and validation set (40 cases) FFPE samples were also selected from a screen of over a thousand patients through an IRB-approved retrospective study at Emory University of men who had undergone radical prostatectomy between 1990 and 1994. Those who were included met specific inclusion criteria, had available tissue specimens, documented long term follow-up and consented to participate or were included by IRB waiver. The cases were assigned prostate ID numbers to protect their identities. These patients did not receive neo-adjuvant or concomitant hormonal therapy. Their demographic, treatment and long-term clinical outcome data have been collected and recorded in an electronic database. Clinical data recorded include PSA measurements, radiological studies and findings, clinical findings, tissue biopsies and additional therapies that the subjects may have received. Clinical data associated with the samples used in this study are given in Supplementary Table S1 (see Supplemental Table S1 at http://ajp.amjpathol.org).

RNA Preparation

Tissue cores (1 mm) were used for RNA preparation rather than sections because of the heterogeneity of samples and the opportunity for obtaining cores with very high percentage tumor content. H&E stained slides were reviewed by a board certified urologic pathologist (AOO) to identify regions of cancer to select corresponding areas for cutting of cores from paraffin blocks. Total RNA was prepared at the Emory Biomarker Service Center from FFPE cores as previously described [4], using the Ambion Recoverall MagMax methodology in 96-well format on a MagMax 96 Liquid Handler Robot (Life Technologies, Carlsbad, CA). FFPE RNA was quantitated quantitated using an Nanodrop spectrophotometer (Wilmington, DE), and tested for RNA integrity and quality by Taqman analysis of the RPL13a ribosomal protein on a HT7900 real-time PCR instrument (Applied Biosystems, Foster City, CA). Samples with sufficient yield (>500 ng), A_{260}/A_{280} ratio > 1.8 and RPL13a C_T values less than 30 cycles were used for miRNA and DASL profiling.

Custom Prostate Cancer DASL Assay Pool (DAP)

The DASL assay enables quantitation of expression of up to 1,536 probes using RNA isolated from archived FFPE tumor tissue samples in a high throughput format [2, 3]. Data from multiple publicly available gene expression datasets [5-8], along with genes involved in prostate cancer progression based on current understanding of the disease [6, 9], were distilled to develop a highly predictive set of 522 genes for use in the DASL assay. Due to specific probe design considerations, this panel had three probes for 497 genes, two probes for 20 genes, and a single probe for five genes, two of which were specific to TMPRSS2-ERG and TMPRSS2-ETV1 fusions transcripts. The unique combination of genes was optimized for performance in the DASL assay using stringent criteria that predicts excellent performance of the primer sets. The panel includes genes found to be correlated with Gleason score in *Liu et al* [10], *Bibikova et al* [11], *True et al* [12], *LaPointe et al* [7], and/or *Singh et al* [13].

6

It also includes prognostic markers from *Dhanasekaran et al* [5] and *Yu et al* [14], and genes associated with metastasis in *Varambally et al* [6]. In addition, a number of genes known from other studies to be critical in prostate cancer such as *NKX3.1, PTEN*, and the androgen receptor are all included in the panel. Other genes that play important roles in the Wnt, Hedgehog, TGFβ, Notch, MAPK and PI3K pathways are also present in this gene set. Finally, primer sets that detect chromosomal translocations in *ERG* [9], *ETV1* [15], and *ETV4* [16] are also included in this panel. The custom prostate cancer panel list of 522 candidate genes (see Supplemental Table S2 at http://ajp.amjpathol.org) was submitted to Illumina for synthesis. The optimal oligonucleotide sequence for each of the 1,536 gene probes was determined using an oligonucleotide scoring algorithm. The oligonucleotide pool or DASL Assay Pool (DAP) was synthesized by Illumina for use with the 96-well Universal Array Matrix (UAM).

The DASL (c<u>D</u>NA-mediated <u>A</u>nnealing, <u>S</u>election, extension and <u>L</u>igation) assay

The DASL assay was performed with our 522-gene custom designed human prostate cancer panel using 200 ng of input RNA at the Emory Biomarker Service Center, Emory University according to the manufacturer's protocols. Samples, including technical replicates (2, 3 or 4) were hybridized on UAMs, and scanned using the BeadStation 500 Instrument (Illumina Inc.). For miRNA DASL assays, the human miRNA v2 DASL panel (Illumina, Inc.), which allows for the determination of expression of 1,146 human miRNAs (> 97% coverage of miRBase release 12) was used. These data are available at GEO under accession number GSE26367.

Data Analysis

DASL fluorescent intensities were interpreted in GenomeStudio, quantile normalized, and exported for meta-analysis. Average signal intensity, genes detected (p-value = 0.01), background, and

noise (standard deviation of background) were analyzed for trends by plate, row, and column. The two endpoints of interest were postoperative biochemical recurrence, defined as two detectable PSA readings (>0.2 ng/ml), and clinical recurrence, defined as evidence of local or metastatic disease. The primary outcome of interest was time to biochemical recurrence following surgery. A local recurrence was defined as recurrence of cancer in the prostatic bed that was detected by either a palpable nodule on digital rectal examination (DRE) and subsequently verified by a positive biopsy, and/or a positive imaging study (Prostascint or CT scan) accompanied by a detectable postoperative PSA result and lack of evidence for metastases. Also, patients whose PSA level decreased following adjuvant pelvic radiation therapy for elevated postoperative PSA were considered as local recurrence cases. A recurrence with metastases was defined as a positive imaging study indicating presence of a tumor outside of the prostatic bed.

To identify important biomarkers and build and evaluate prediction models for prostate cancer recurrence, we adopted the following strategy. In the training step, the prediction model was built based on the time to biochemical recurrence. Specifically, we first fit a univariate Cox proportional hazard (PH) model for each individual oligonucleotide probe using the training data set, and a set of important mRNA and miRNA probes were then preselected based on a false discovery rate (FDR) threshold of 0.30. Next, to identify the optimal prediction score based on the preselected probes, we fit a lasso Cox PH model [17, 18] using the training data set, where the tuning parameter for lasso was selected using a leave-one-out cross-validation technique [18]. The lasso Cox PH model was fitted first using the set of preselected mRNA probes only and then using the complete set of preselected mRNA and miRNA probes, resulting in an optimal mRNA panel and an optimal combined mRNA/miRNA panel, respectively. Based on each biomarker panel, a final prediction model for recurrence was built to also incorporate relevant clinical biomarkers, namely, T-stage, PSA and Gleason score, through

fitting Cox PH models. For comparison, we also built a prediction model using only clinical information, namely, T-stage, PSA and Gleason score, through fitting a Cox PH model.

To evaluate and validate the final prediction models obtained from the training phase, 79 samples from 40 patients were used and replicate samples from the same patient were again averaged to generate a single average signal for each patient. Each prediction model from the training phase was used to generate a predictive score for each subject in the validation data set, and subjects were subsequently divided into high and low scoring groups using the median predictive score. Kaplan Meier analysis was performed to compare the time to biochemical recurrence, between high (poor score) and low (good score) risk groups, and the statistical significance was determined using the log-rank test. Similarly, we also evaluated the final model that uses the combined mRNA/miRNA panel for predicting time to clinical recurence in both training and validation data sets.

Missing data are present in this study, in particular, for clinical recurrence, PSA and T-stage data. We adopted the available-case approach[19] in our analyses and the sample sizes used in each step of building and evaluating prediction models may be less than the total sample size.

RESULTS

Custom Prostate DASL profiling

We performed DASL expression profiling with our custom-designed prostate cancer panel (see Materials and Methods section) and the Illumina DASL microRNA (miRNA) panel on 70 prostatectomy patient samples to identify biomarkers predictive of recurrence. An independent validation profiling experiment was performed on 40 additional samples. MiRNA probes were filtered to retain only those that were present on the miRNA microarrays used for both the training and validation sets, reducing the total number of probes examined to 403 miRNA probes. The training set included 29 cases with observed biochemical PSA recurrence (median time to recurrence = 19 months), and 41 cases censored, i.e., without observed recurrence during follow-up (median follow-up time = 83.0 months). A summary of the clinical characteristics of the training and validation sets of samples is provided in **Table 1**. The complete dataset for the combined mRNA and miRNA data are provided in Supplementary Table S3 for the training set and Supplementary Table S4 for the validation set (see http://ajp.amjpathol.org).

Integrated DASL biomarker analysis

After fitting a univariate Cox proportional hazard (PH) model for each individual probe using the training data, a set of 27 important probes were preselected based on an FDR threshold of 0.30 (see Supplementary Table S5 at http://ajp.amjpathol.org). Next, to identify the optimal prediction score based on the preselected probes, we fit a lasso Cox proportional hazard (PH) model[17, 18] first using the set of 25 preselected mRNA probes only, resulting in a panel of nine protein-coding genes shown in **Table 2** (*RAD23B, FBP1, TNFRSF1A, NOTCH3, ETV1, BID, SIM2, ANXA1*, and *BCL2*). A final prediction model was then built to include the predictive score based on this panel of nine mRNA

biomarkers as well as the relevant clinical biomarkers including T-stage, PSA and Gleason score, which could be used to predict recurrence following radical prostatectomy. Kaplan-Meier analysis (**Figure 1A**) demonstrated that these probes could significantly discriminate patients at higher and lower risk of recurrence by the log rank test ($p < 0.001$). We next applied the final predictive model developed on the training set to the validation set, a separate, independent DASL profiling experiment performed on a different day. Kaplan-Meier analysis (**Figure 1B**) on this validation set determined that the model could discriminate patients at higher and lower risk of recurrence ($p = 0.010$).

Subsequently, we repeated the above training procedure using the complete set of 27 preselected mRNA and miRNA probes, and we identified an optimal panel of ten mRNAs and two microRNAs (**Table 3**) and built a final prediction model for prostate cancer biochemical recurrence, which again included relevant clinical biomarkers. Kaplan-Meier analysis and the log-rank test determined that this panel could also significantly discriminate patients at higher and lower risk of recurrence both in the training set ($p < 0.001$, **Figure 1C**) and in the validation set ($p = 0.013$, **Figure 1D**).

Prediction of Cases with a Gleason Score 7

Prediction of recurrence for patients with a Gleason score 7 is particularly difficult. In order to address this issue, we applied the biomarker panels to the subset of cases in the training and validation sets that had a Gleason score 7. The prediction model based on the nine-mRNA panel was significant at discriminating biochemical recurrence in Gleason score 7 cases in both the training set ($p < 0.001$, Figure 2A) and the validation set ($p = 0.027$, Figure 2B). For the prediciton model based on the combined panel of ten mRNAs and two miRNAs in Table 3, the predictive value was again significant

for both the training set (p = < 0.001, Figure 2C) and the validation set (p = 0.010, Figure 2D). A

summary of the p-values for predicting biochemical recurrence is given in **Table 4**. In all cases, the

prediction models that use one of the two gene biomarker panels plus clinical information outperforms

the prediction model using only clinical information.

Analysis of clinical recurrence

Although most patients who have clinical recurrence following prostatectomy also have

biochemical recurrence, there is a significant population of patients with biochemical recurrence who

do not have clinically significant recurrences observed during their follow-ups. To evaluate our

biomarker panel of biochemical recurrence for predicting the clinical recurrence, we tested the

prediction model based on the combined mRNA/miRNA panel in the same training and validation

samples using their clinical recurrence outcome data. Unfortunately, clinical recurrence data was

lacking on some of the samples, and the total number of samples used in the training set was reduced.

In the training data, the combined mRNA/miRNA panel was highly significant for predicting clinical

recurrence in all patients (p=0.002) as well as in the subset of patients with a Gleason score 7

(p=0.004); in the validation data, it was also significant for predicting recurrence in patients with a

Gleason score 7 (p=0.023) and trended towards significance in all patients (p=0.078). A summary of

the p-values for predicting clinical recurrence is given in **Table 5**. In all cases, the prediction model

that uses the combined mRNA and miRNA panel plus the clinical information, again, outperforms the

prediction model that uses only the clinical information.

We also performed an analysis to construct a predictive set of biomarkers based on the clinical

recurrence data instead of biochemical recurrence. Only three probes passed the initial preselection

step for the univariate Cox PH modeling, all corresponding to the *ETV1* gene, which is likely due to

the considerably fewer number of clinical recurrences in the training set as well as the smaller total sample size. Furthermore, the prediction model built on this set of gene biomarkers did not perform as well as the models built on biochemical recurrence (data not shown).

DISCUSSION

In the past few years, enormous progress has been made in developing technologies to exploit FFPE tumor tissue samples for gene expression and proteomic analysis. The use of FFPE tissues as a starting material is attractive because this approach should make biomarkers identified in this way much easier to translate into widespread clinical practice. DASL profiling makes it possible to define gene sets using FFPE prostate cancer tissues that could have potential prognostic and predictive value. For example, the DASL assay has been used recently to identify a 16-gene set that correlates with prostate cancer relapse [11]. There was no overlap between our panel of ten mRNA and two miRNA biomarkers described here and the previously described 16-gene panel even though ten of the genes in the 16-gene panel previously reported were included in our 522 custom prostate DASL panel. When we analyzed the performance of the probes corresponding to those ten mRNAs in our dataset, we found that they were not able to significantly discriminate patients at higher or lower risk of recurrence. In this previous study, the gene signature selection and prediction model building were performed in separate steps and the signature selection was based on the correlation between the gene expression and Gleason score rather than between the gene expression and time to biochemical recurrence; our analytic approach overcomes these limitations. Specifically, our approach of building (training) prediction models takes advantage of recent advancement in regularized regression models for survival outcomes[17, 18]; regularized regression models can achieve simultaneous feature selection and model estimation and avoid model overfitting, leading to better prediction performance. Our use of a pre-selection step is similar to the recently proposed sure independence screening methods[20, 21], which have been shown to achieve better performance in the presence of high-dimensional data for survival analysis compared to regularized regression without a pre-selection step[22].

Two other recent studies have employed DASL profiling to prostate cancer, but not detected

any signature that improved upon clinical models in validation sets[23, 24]. While these studies used large cohorts with long-term follow-up, they examined different panels of mRNA transcripts and did not include probes corresponding to miRNA genes. Moreover, these earlier studies suggested that tumor heterogeneity may play an important role in confounding signature identification. For our study of prostatectomy specimens, we identified the most prominent tumor lesion, and used a tissue core sample from that region to minimize stromal contributions and tumor heterogeneity.

In our twelve-gene predictive biomarker panel, nine of the genes are positively associated with recurrence, and three are negatively associated with recurrence. The nine genes positively associated with recurrence included *miR-519d*, Notch homolog 3 (Notch3), Fructose-1,6-bisphosphatase 1 (FBP1), ETS variant gene 1 (ETV1), BH3 interacting domain death agonist (BID), Single-Minded homolog 2 (SIM2), RAD23 homolog B (RAD23B), LETM1 domain containing 1 (LETMD1), and Cyclin G2 (CCNG2). Little is known about *miR-519d* other than it may be associated with obesity[25]. NOTCH3 is one of four Notch family receptors in humans, and Notch signaling has been shown to be important for prostate cancer cell growth, migration, and invasion[26, 27] as well as normal prostate development[28, 29]. FBP1 is expressed in the prostate and is involved in gluconeogenesis[30]. The identification of this metabolic enzyme as a biomarker of recurrence is initially surprising, but given the recent identification of isocitrate dehydrogenase 1 (IDH1) mutations in glioblastoma[31], and the fact that FBP1 was overexpressed in independent microarray analyses of prostate cancers[7, 32], the potential of FBP1 as a biomarker should not be underestimated. ETV1 is well established as one of the commonly recurrent translocations found in prostate cancers[9, 15], and has been used in clinical models of recurrence following prostatectomy[33]. BID is a pro-apoptotic protein that binds to BCL2 and potentiates apoptotic responses upon cleavage in response to tumor necrosis factor alpha (TNFα) and other death receptors[34, 35]. SIM2 was identified as a potential biomarker of prostate cancer in 2002[36]

and later independently confirmed by *Halvorsen et al*[37] and *Arredouani et al*[38]. SIM2 functions as a transcription factor that represses the proapoptotic gene BNIP3[39]. RAD23B plays a critical role in DNA damage recognition and nucleotide excision repair[40], as well as inhibiting MDM2-mediated degradation of the p53 tumor suppressor[41]. LETMD1 (also known as HCCR) is an oncogene that is induced by Wnt[42] and PI3K/AKT signaling[43], inhibits p53 function[44], and is a biomarker for hepatocellular[45] and breast[46] cancers. Cyclin G2 is an atypical cyclin that is induced by DNA damage[47] in a p53-independent manner, as well as by PI3K/AKT/FOXO signals[48], and induces p53-dependent cell cycle arrest[49].

The three genes in the predictive biomarker panel negatively associated with recurrence were *miR-647*, the TNFα receptor (TNFRSF1A), and annexin A1 (ANXA1). While little is known about *miR-647*, TNFRSF1A (also known as TNFR1) mediates pro-apoptotic responses to TNFα ligand[50, 51]. Annexin A1 expression is reduced in early onset prostate cancer[52] and high-grade prostatic intraepithelial neoplasia[53]. ANXA1 plays important roles in vesicle trafficking and reduced ANXA1 promotes EMT and metastasis [54], and upregulates autocrine IL-6 signaling[55]. Thus, as a whole, this panel of biomarkers appears to reflect changes in DNA stability, PI3K signaling, p53 activity, apoptosis, and differentiation consistent with more aggressive disease.

Although this study goes beyond a pilot study, enhanced by selection of samples from multiple institutions, the number of specimens tested is still relatively small. Re-analysis of our data using only the Emory samples for the training set did not identify any significant probes, likely due to the substantially smaller sample size. Thus, while the performance of our panel of biomarkers is significant, even for Gleason score 7 patients, future studies beyond the scope of this work will be necessary to perform independent validation on much larger sample sets with greater statistical power. Moreover, it is now feasible to perform DASL assays on virtually the entire genome, in an assay that

queries 24,526 transcripts derived from the RefSeq database. Future studies will test combined mRNA

and miRNA biomarker panels, and query the entire genome to determine if other biomarker panels can

achieve even greater success in prediction of biochemical and clinical recurrence of prostate cancer.

Planned larger scale validation studies will determine whether these biomarkers are predictive for

Gleason score 7 cases, and their utility at predicting clinical as well as biochemical recurrence.

REFERENCES

1.	ACS: Cancer Facts and Figures 2009. Edited by Atlanta, American Cancer Society, 2009, p.
2.	Bibikova M, Talantov D, Chudin E, Yeakley JM, Chen J, Doucet D, Wickham E, Atkins D, Barker D, Chee M, Wang Y, Fan JB: Quantitative gene expression profiling in formalin-fixed, paraffin-embedded tissues using universal bead arrays, Am J Pathol 2004, 165:1799-1807
3.	Fan JB, Yeakley JM, Bibikova M, Chudin E, Wickham E, Chen J, Doucet D, Rigault P, Zhang B, Shen R, McBride C, Li HR, Fu XD, Oliphant A, Barker DL, Chee MS: A versatile assay for high-throughput gene expression profiling on universal array matrices, Genome Res 2004, 14:878-885
4.	Abramovitz M, Ordanic-Kodani M, Wang Y, Li Z, Catzavelos C, Bouzyk M, Sledge GW, Jr., Moreno CS, Leyland-Jones B: Optimization of RNA extraction from FFPE tissues for expression profiling in the DASL assay, Biotechniques 2008, 44:417-423
5.	Dhanasekaran SM, Barrette TR, Ghosh D, Shah R, Varambally S, Kurachi K, Pienta KJ, Rubin MA, Chinnaiyan AM: Delineation of prognostic biomarkers in prostate cancer, Nature 2001, 412:822-826.
6.	Varambally S, Yu J, Laxman B, Rhodes DR, Mehra R, Tomlins SA, Shah RB, Chandran U, Monzon FA, Becich MJ, Wei JT, Pienta KJ, Ghosh D, Rubin MA, Chinnaiyan AM: Integrative genomic and proteomic analysis of prostate cancer reveals signatures of metastatic progression, Cancer Cell 2005, 8:393-406
7.	Lapointe J, Li C, Higgins JP, Van De Rijn M, Bair E, Montgomery K, Ferrari M, Egevad L, Rayford W, Bergerheim U, Ekman P, DeMarzo AM, Tibshirani R, Botstein D, Brown PO, Brooks JD, Pollack JR: Gene expression profiling identifies clinically relevant subtypes of prostate cancer, Proc Natl Acad Sci U S A 2004, 101:811-816
8.	LaTulippe E, Satagopan J, Smith A, Scher H, Scardino P, Reuter V, Gerald WL: Comprehensive gene expression analysis of prostate cancer reveals distinct transcriptional programs associated with metastatic disease, Cancer Res 2002, 62:4499-4506
9.	Tomlins SA, Rhodes DR, Perner S, Dhanasekaran SM, Mehra R, Sun XW, Varambally S, Cao X, Tchinda J, Kuefer R, Lee C, Montie JE, Shah RB, Pienta KJ, Rubin MA, Chinnaiyan AM: Recurrent fusion of TMPRSS2 and ETS transcription factor genes in prostate cancer, Science 2005, 310:644-648
10.	Liu P, Ramachandran S, Ali-Seyed M, Scharer CD, Laycock N, Dalton WB, Williams H, Karanam S, Datta MW, Jaye D, Moreno CS: Sex-Determining Region Y Box 4 is a Transforming Oncogene in Human Prostate Cancer Cells, Cancer Res 2006, 46:4011-4019
11.	Bibikova M, Chudin E, Arsanjani A, Zhou L, Garcia EW, Modder J, Kostelec M, Barker D, Downs T, Fan JB, Wang-Rodriguez J: Expression signatures that correlated with Gleason score and relapse in prostate cancer, Genomics 2007, 89:666-672
12.	True L, Coleman I, Hawley S, Huang CY, Gifford D, Coleman R, Beer TM, Gelmann E, Datta M, Mostaghel E, Knudsen B, Lange P, Vessella R, Lin D, Hood L, Nelson PS: A molecular correlate to the Gleason grading system for prostate adenocarcinoma, Proc Natl Acad Sci U S A 2006, 103:10991-10996
13.	Singh D, Febbo PG, Ross K, Jackson DG, Manola J, Ladd C, Tamayo P, Renshaw AA, D'Amico AV, Richie JP, Lander ES, Loda M, Kantoff PW, Golub TR, Sellers WR: Gene expression correlates of clinical prostate cancer behavior, Cancer Cell 2002, 1: 203-209

14.   Yu YP, Landsittel D, Jing L, Nelson J, Ren B, Liu L, McDonald C, Thomas R, Dhir R, Finkelstein S, Michalopoulos G, Becich M, Luo JH: Gene expression alterations in prostate cancer predicting tumor aggression and preceding development of malignancy, J Clin Oncol 2004, 22:2790-2799

15.   Tomlins SA, Laxman B, Dhanasekaran SM, Helgeson BE, Cao X, Morris DS, Menon A, Jing X, Cao Q, Han B, Yu J, Wang L, Montie JE, Rubin MA, Pienta KJ, Roulston D, Shah RB, Varambally S, Mehra R, Chinnaiyan AM: Distinct classes of chromosomal rearrangements create oncogenic ETS gene fusions in prostate cancer, Nature 2007, 448:595-599

16.   Tomlins SA, Mehra R, Rhodes DR, Smith LR, Roulston D, Helgeson BE, Cao X, Wei JT, Rubin MA, Shah RB, Chinnaiyan AM: TMPRSS2:ETV4 gene fusions define a third molecular subtype of prostate cancer, Cancer Res 2006, 66:3396-3400

17.   Tibshirani R: The lasso method for variable selection in the Cox model, Stat Med 1997, 16:385-395

18.   Goeman JJ: L1 penalized estimation in the Cox proportional hazards model, Biom J 2010, 52:70-84

19.   Little RJA, Rubin DB: Statistical Analysis with Missing Data. Edited by New York, NY, Wiley, 2002,

20.   Fan J, Ly J: Sure Independence Screening for Ultrahigh Dimensional Feature Space, Journal of the Royal Statistical Society 2008, 70:849-911

21.   Fan J, Song R: Sure Independence Screening in Generalized Linear Models and NP-Dimensionality. , The Annals of Statistics 2010, in press

22.   Zhao SD, Li Y: Principled sure independence screening for Cox models with ultra-high-dimensional covariates, Harvard University Biostatistics Working Paper Series 2010,

23.   Sboner A, Demichelis F, Calza S, Pawitan Y, Setlur SR, Hoshida Y, Perner S, Adami HO, Fall K, Mucci LA, Kantoff PW, Stampfer M, Andersson SO, Varenhorst E, Johansson JE, Gerstein MB, Golub TR, Rubin MA, Andren O: Molecular sampling of prostate cancer: a dilemma for predicting disease progression, BMC Med Genomics 2010, 3:8

24.   Nakagawa T, Kollmeyer TM, Morlan BW, Anderson SK, Bergstralh EJ, Davis BJ, Asmann YW, Klee GG, Ballman KV, Jenkins RB: A tissue biomarker panel predicting systemic progression after PSA recurrence post-definitive prostate cancer therapy, PLoS ONE 2008, 3:e2318

25.   Martinelli R, Nardelli C, Pilone V, Buonomo T, Liguori R, Castano I, Buono P, Masone S, Persico G, Forestieri P, Pastore L, Sacchetti L: miR-519d Overexpression Is Associated With Human Obesity, Obesity (Silver Spring) 2010,

26.   Wang Z, Li Y, Ahmad A, Banerjee S, Azmi AS, Kong D, Wojewoda C, Miele L, Sarkar FH: Down-regulation of Notch-1 is associated with Akt and FoxM1 in inducing cell growth inhibition and apoptosis in prostate cancer cells, J Cell Biochem 2010,

27.   Wang Z, Li Y, Banerjee S, Kong D, Ahmad A, Nogueira V, Hay N, Sarkar FH: Down-regulation of Notch-1 and Jagged-1 inhibits prostate cancer cell growth, migration and invasion, and induces apoptosis via inactivation of Akt, mTOR, and NF-kappaB signaling pathways, J Cell Biochem 2010, 109:726-736

28.   Leong KG, Gao WQ: The Notch pathway in prostate development and cancer, Differentiation 2008, 76:699-716

29.   Wang XD, Leow CC, Zha J, Tang Z, Modrusan Z, Radtke F, Aguet M, de Sauvage FJ, Gao WQ: Notch signaling is required for normal prostatic epithelial cell proliferation and differentiation, Dev Biol 2006, 290:66-80

30.   Yanez AJ, Nualart F, Droppelmann C, Bertinat R, Brito M, Concha, II, Slebe JC: Broad expression of fructose-1,6-bisphosphatase and phosphoenolpyruvate carboxykinase provide evidence for gluconeogenesis in human tissues other than liver and kidney, J Cell Physiol 2003, 197:189-197

31.   Parsons DW, Jones S, Zhang X, Lin JC, Leary RJ, Angenendt P, Mankoo P, Carter H, Siu IM, Gallia GL, Olivi A, McLendon R, Rasheed BA, Keir S, Nikolskaya T, Nikolsky Y, Busam DA, Tekleab H, Diaz LA, Jr., Hartigan J, Smith DR, Strausberg RL, Marie SK, Shinjo SM, Yan H, Riggins GJ, Bigner DD, Karchin R, Papadopoulos N, Parmigiani G, Vogelstein B, Velculescu VE, Kinzler KW: An integrated genomic analysis of human glioblastoma multiforme, Science 2008, 321:1807-1812

32.   Liu P, Ramachandran S, Ali Seyed M, Scharer CD, Laycock N, Dalton WB, Williams H, Karanam S, Datta MW, Jaye DL, Moreno CS: Sex-determining region Y box 4 is a transforming oncogene in human prostate cancer cells, Cancer Res 2006, 66:4011-4019

33.   Cheville JC, Karnes RJ, Therneau TM, Kosari F, Munz JM, Tillmans L, Basal E, Rangel LJ, Bergstralh E, Kovtun IV, Savci-Heijink CD, Klee EW, Vasmatzis G: Gene panel model predictive of outcome in men at high-risk of systemic progression and death from prostate cancer after radical retropubic prostatectomy, J Clin Oncol 2008, 26:3930-3936

34.   Kulik G, Carson JP, Vomastek T, Overman K, Gooch BD, Srinivasula S, Alnemri E, Nunez G, Weber MJ: Tumor necrosis factor alpha induces BID cleavage and bypasses antiapoptotic signals in prostate cancer LNCaP cells, Cancer Res 2001, 61:2713-2719

35.   Luo X, Budihardjo I, Zou H, Slaughter C, Wang X: Bid, a Bcl2 interacting protein, mediates cytochrome c release from mitochondria in response to activation of cell surface death receptors, Cell 1998, 94:481-490

36.   Deyoung MP, Scheurle D, Damania H, Zylberberg C, Narayanan R: Down's syndrome-associated single minded gene as a novel tumor marker, Anticancer Res 2002, 22:3149-3157

37.   Halvorsen OJ, Rostad K, Oyan AM, Puntervoll H, Bo TH, Stordrange L, Olsen S, Haukaas SA, Hood L, Jonassen I, Kalland KH, Akslen LA: Increased expression of SIM2-s protein is a novel marker of aggressive prostate cancer, Clin Cancer Res 2007, 13:892-897

38.   Arredouani MS, Lu B, Bhasin M, Eljanne M, Yue W, Mosquera JM, Bubley GJ, Li V, Rubin MA, Libermann TA, Sanda MG: Identification of the transcription factor single-minded homologue 2 as a potential biomarker and immunotherapy target in prostate cancer, Clin Cancer Res 2009, 15:5794-5802

39.   Farrall AL, Whitelaw ML: The HIF1alpha-inducible pro-cell death gene BNIP3 is a novel target of SIM2s repression through cross-talk on the hypoxia response element, Oncogene 2009, 28:3671-3680

40.   Thoma BS, Vasquez KM: Critical DNA damage recognition functions of XPC-hHR23B and XPA-RPA in nucleotide excision repair, Mol Carcinog 2003, 38:1-13

41.   Brignone C, Bradley KE, Kisselev AF, Grossman SR: A post-ubiquitination role for MDM2 and hHR23A in the p53 degradation pathway, Oncogene 2004, 23:4121-4129

42.   Cho GW, Kim MH, Kim SH, Ha SA, Kim HK, Kim S, Kim JW: TCF/beta-catenin plays an important role in HCCR-1 oncogene expression, BMC Mol Biol 2009, 10:42

43.   Cho GW, Shin SM, Namkoong H, Kim HK, Ha SA, Hur SY, Kim TE, Chai YG, Kim JW: The phosphatidylinositol 3-kinase/Akt pathway regulates the HCCR-1 oncogene expression, Gene 2006, 384:18-26

44.   Ko J, Lee YH, Hwang SY, Lee YS, Shin SM, Hwang JH, Kim J, Kim YW, Jang SW, Ryoo ZY, Kim IK, Namkoong SE, Kim JW: Identification and differential expression of novel human

cervical cancer oncogene HCCR-2 in human cancers and its involvement in p53 stabilization, Oncogene 2003, 22:4679-4689

45.   Yoon SK, Lim NK, Ha SA, Park YG, Choi JY, Chung KW, Sun HS, Choi MJ, Chung J, Wands JR, Kim JW: The human cervical cancer oncogene protein is a biomarker for human hepatocellular carcinoma, Cancer Res 2004, 64:5434-5441

46.   Jung SS, Park HS, Lee IJ, Namkoong H, Shin SM, Cho GW, Ha SA, Park YG, Lee YS, Ko J, Kim JW: The HCCR oncoprotein as a biomarker for human breast cancer, Clin Cancer Res 2005, 11:7700-7708

47.   Bates S, Rowan S, Vousden KH: Characterisation of human cyclin G1 and G2: DNA damage inducible genes, Oncogene 1996, 13:1103-1109

48.   Martinez-Gac L, Marques M, Garcia Z, Campanero MR, Carrera AC: Control of cyclin G2 mRNA expression by forkhead transcription factors: novel mechanism for cell cycle control by phosphoinositide 3-kinase and forkhead, Mol Cell Biol 2004, 24:2181-2189

49.   Arachchige Don AS, Dallapiazza RF, Bennin DA, Brake T, Cowan CE, Horne MC: Cyclin G2 is a centrosome-associated nucleocytoplasmic shuttling protein that influences microtubule stability and induces a p53-dependent cell cycle arrest, Exp Cell Res 2006, 312:4181-4204

50.   Hsu H, Xiong J, Goeddel DV: The TNF receptor 1-associated protein TRADD signals cell death and NF-kappa B activation, Cell 1995, 81:495-504

51.   Thorburn A: Death receptor-induced cell killing, Cell Signal 2004, 16:139-144

52.   Paweletz CP, Ornstein DK, Roth MJ, Bichsel VE, Gillespie JW, Calvert VS, Vocke CD, Hewitt SM, Duray PH, Herring J, Wang QH, Hu N, Linehan WM, Taylor PR, Liotta LA, Emmert-Buck MR, Petricoin EF, 3rd: Loss of annexin 1 correlates with early onset of tumorigenesis in esophageal and prostate carcinoma, Cancer Res 2000, 60:6293-6297

53.   Patton KT, Chen HM, Joseph L, Yang XJ: Decreased annexin I expression in prostatic adenocarcinoma and in high-grade prostatic intraepithelial neoplasia, Histopathology 2005, 47:597-601

54.   Maschler S, Gebeshuber CA, Wiedemann EM, Alacakaptan M, Schreiber M, Custic I, Beug H: Annexin A1 attenuates EMT and metastatic potential in breast cancer, EMBO Mol Med 2010,

55.   Inokuchi J, Lau A, Tyson DR, Ornstein DK: Loss of annexin A1 disrupts normal prostate glandular structure by inducing autocrine IL-6 signaling, Carcinogenesis 2009, 30:1082-1088

56.   DeYoung MP, Tress M, Narayanan R: Identification of Down's syndrome critical locus gene SIM2-s as a drug therapy target for solid tumors, Proc Natl Acad Sci U S A 2003, 100:4760-4765

57.   Alnemri ES, Robertson NM, Fernandes TF, Croce CM, Litwack G: Overexpressed full-length human BCL2 extends the survival of baculovirus-infected Sf9 insect cells, Proc Natl Acad Sci U S A 1992, 89:7295-7299

58.   Wagner AJ, Small MB, Hay N: Myc-mediated apoptosis is blocked by ectopic expression of Bcl-2, Mol Cell Biol 1993, 13:2432-2440

ACKNOWLEDGEMENTS

TABLES

	Training Set (Total)	Training Set (No BCR)	Training Set (BCR)	Validation Set (Total)	Validation Set (No BCR)	Validation Set (BCR)
Number Cases	70	41	29	40	27	13
Clinical Recurrence	8	0	8	11	0	11
No Clinical Recurrence	57	41	16	29	27	2
Median Time F/U (months)	84	83	81	74	75	73
Median Time to BCR (months)	19	N/A	19	14	N/A	14
Median Time no BCR (months)	48	83	19	34.5	56	14
Gleason Score (Avg +/- SD)	6.9 +/- 0.6	6.7 +/- 0.6	7.0 +/- 0.6	7.0 +/- 0.8	6.8 +/- 0.7	7.4 +/- 1
PSA (Avg +/- SD)	9.2 +/- 5.4	8.7 +/- 6.4	9.9 +/- 3.8	12.7 +/- 8.4	12.4 +/- 9.9	13.1 +/- 5.3
Age (Avg +/- SD)	61.9 +/- 7.7	61.2 +/- 7.7	62.9 +/- 7.8	63.6 +/- 8.4	63.5 +/- 8.3	64 +/- 8.9

Table 1: A summary of the clinical characteristics of the training and validation sets of patient samples. (BCR = Biochemical Recurrence, F/U = follow up, PSA = prostate specific antigen, SD = standard deviation).

Symbol	Description	Coefficient	References
RAD23B	RAD23 homolog B	0.152	[40, 41]
FBP1	Fructose-1,6-bisphosphatase 1	0.310	[7, 30, 32]
TNFRSF1A	Tumor Necrosis Factor Receptor Superfamily, Member 1A	-0.560	[50, 51]
NOTCH3	Notch homolog 3	0.426	[26, 27]
ETV1	Ets Variant Gene 1 (ETV1)	0.157	[9, 15]
BID	BH3 Interacting Domain Death Agonist (BID)	0.248	[34, 35]
SIM2	Single-Minded Homolog 2	0.043	[36-38, 56]
ANXA1	Annexin A1	-0.185	[52-55]
BCL2	B-cell CLL/lymphoma 2	0.028	[57, 58]

Table 2: Nine-gene predictor of prostate cancer recurrence following surgery. Coefficient is derived from the lasso Cox proportion hazards model and was used for computing the predictive score.

Positive coefficients indicate a positive association with recurrence, and negative coefficients a negative association with recurrence.

Symbol	Description	Coefficient	References
RAD23B	RAD23 homolog B	0.070	[40, 41]
FBP1	Fructose-1,6-bisphosphatase 1	0.251	[7, 30, 32]
TNFRSF1A	Tumor necrosis factor receptor superfamily, member 1A	-0.588	[50, 51]
CCNG2	Cyclin G2	0.008	[47-49]
hsa-miR-647	hsa-miR-647	-0.318	
LETMD1	LETM1 domain containing 1	0.063	[42-44, 46]
NOTCH3	Notch homolog 3	0.367	[26, 27]
ETV1	ETS variant gene 1 (ETV1)	0.179	[9, 15]
hsa-miR-519d	hsa-miR-519d	0.551	[25]
BID	BH3 interacting domain death agonist (BID)	0.128	[34, 35]
SIM2	Single-minded homolog 2	0.124	[36-38, 56]
ANXA1	Annexin A1	-0.143	[52-55]

Table 3: Twelve-gene predictor of prostate cancer recurrence following surgery using ten mRNAs and two microRNAs. Coefficient is derived from the lasso Cox proportion hazards model and was used for computing the predictive score. Positive coefficients indicate a positive association with recurrence, and negative coefficients a negative association with recurrence.

Training Set	mRNA panel	Combined mRNA/miRNA panel	Clinical Information Only
All Cases (n=61)	**<0.001**	**<0.001**	0.096
Gleason score 7 (n=42)	**<0.001**	**<0.001**	0.641
Validation Set	mRNA panel	Combined mRNA/miRNA panel	Clinical Information Only
All Cases (n=35)	0.010	**0.013**	**0.020**
Gleason score 7 (n=19)	0.027	**0.010**	**0.028**

Table 4: Summary of p-values (Logrank test) of prediction of biochemical recurrence on training and validation sets for the entire dataset and the subset of Gleason score 7 cases using two biomarker panels, all of which are adjusted for T-stage, PSA, and Gleason score, or using clinical information only. Significant p-values are indicated in bold.
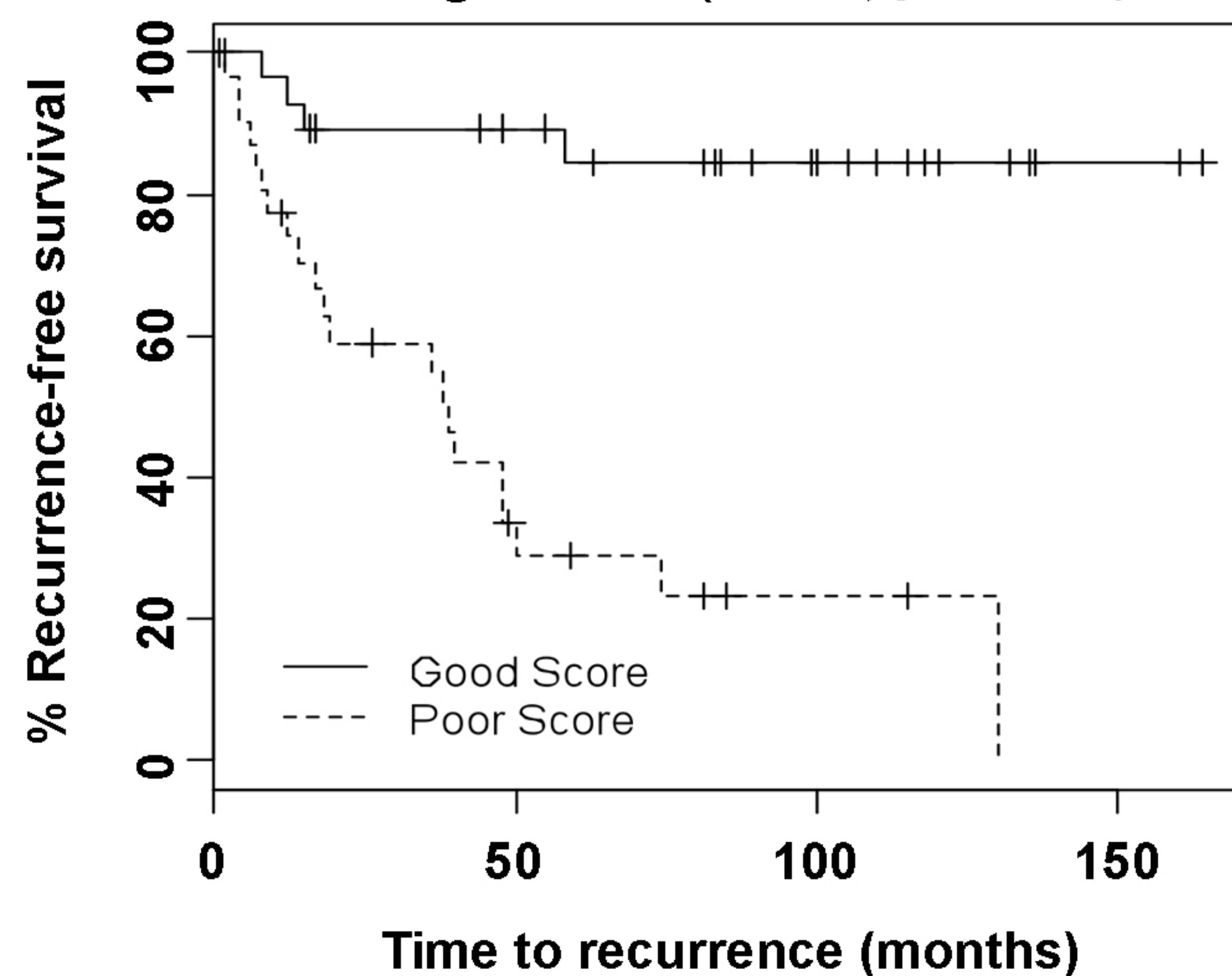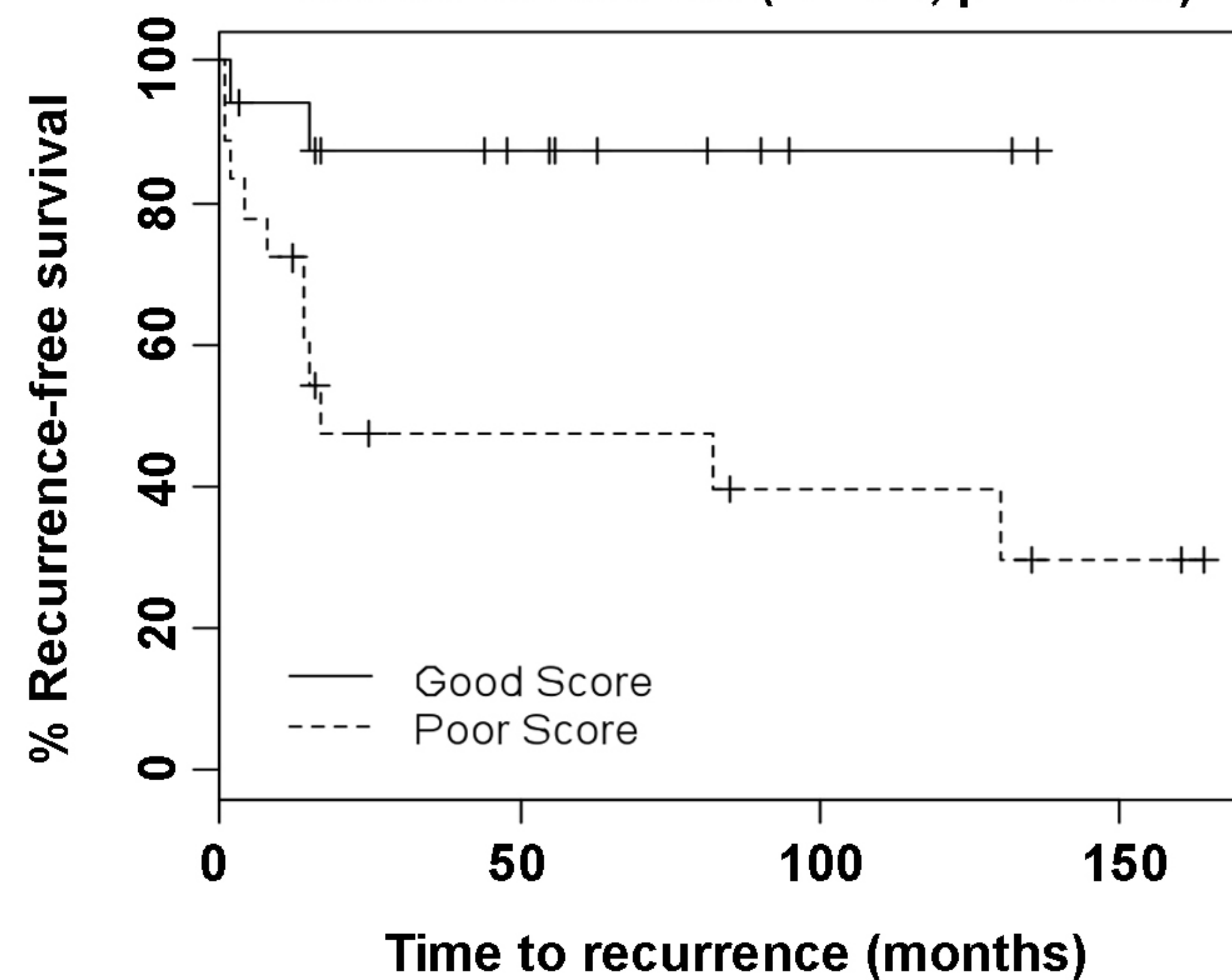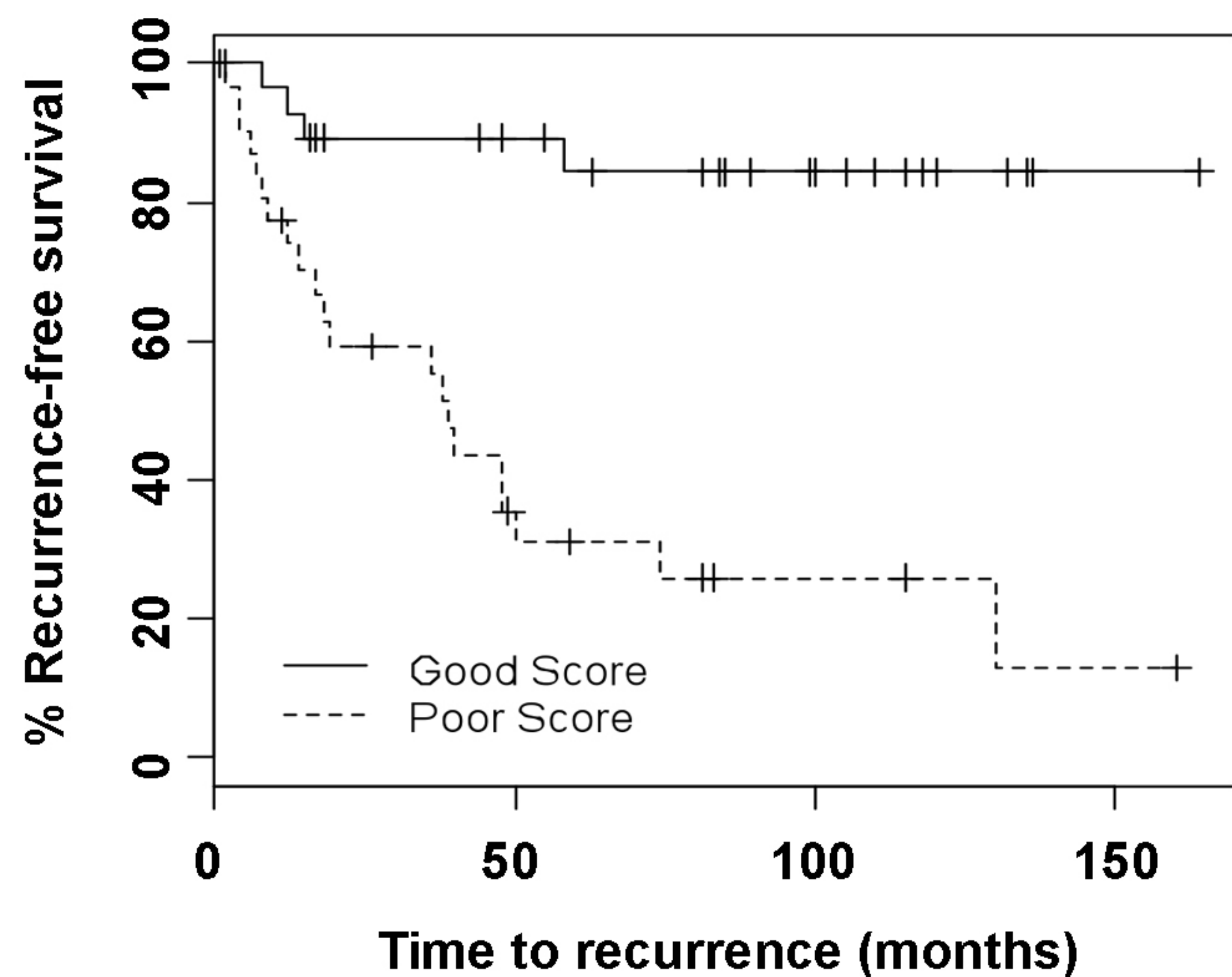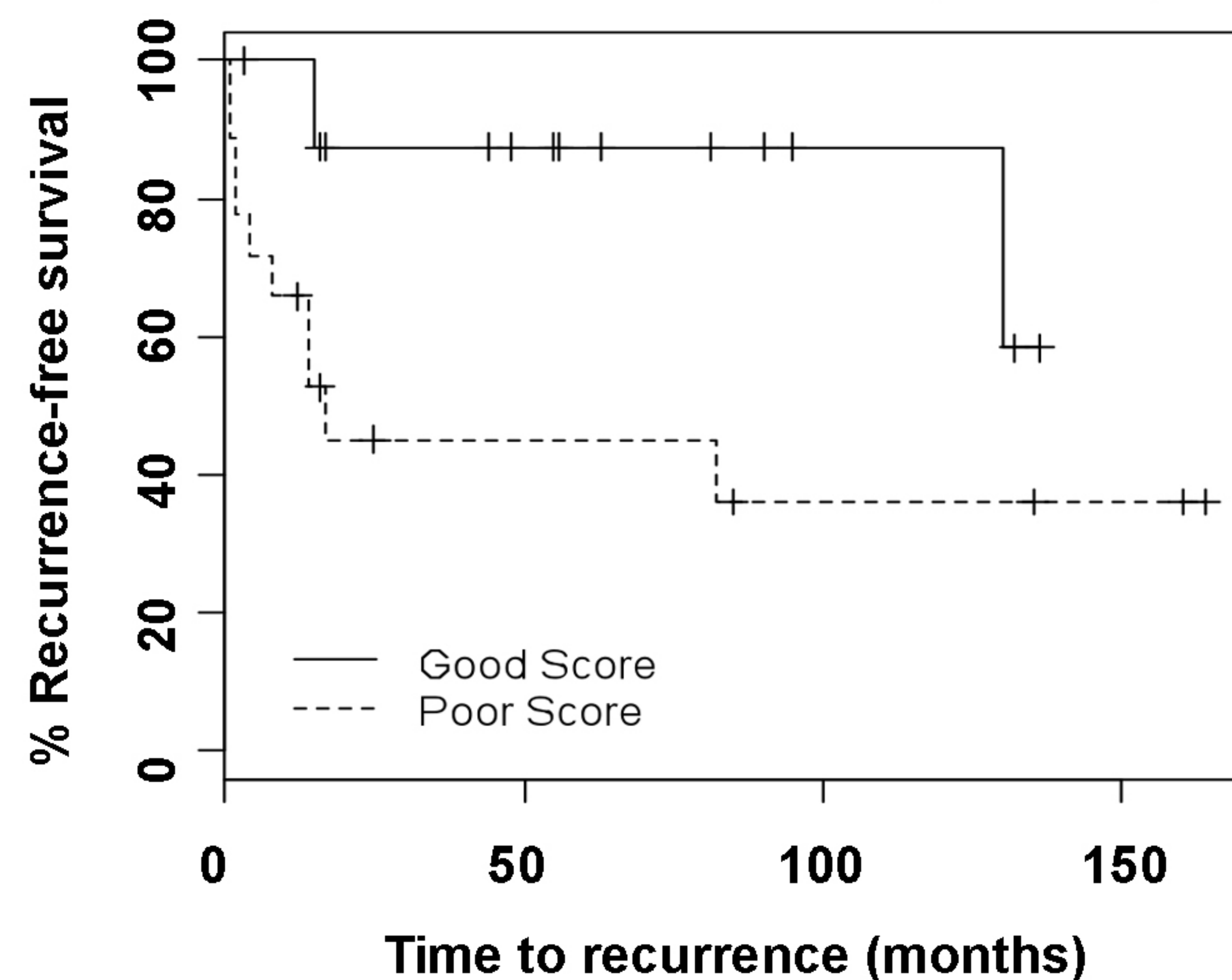
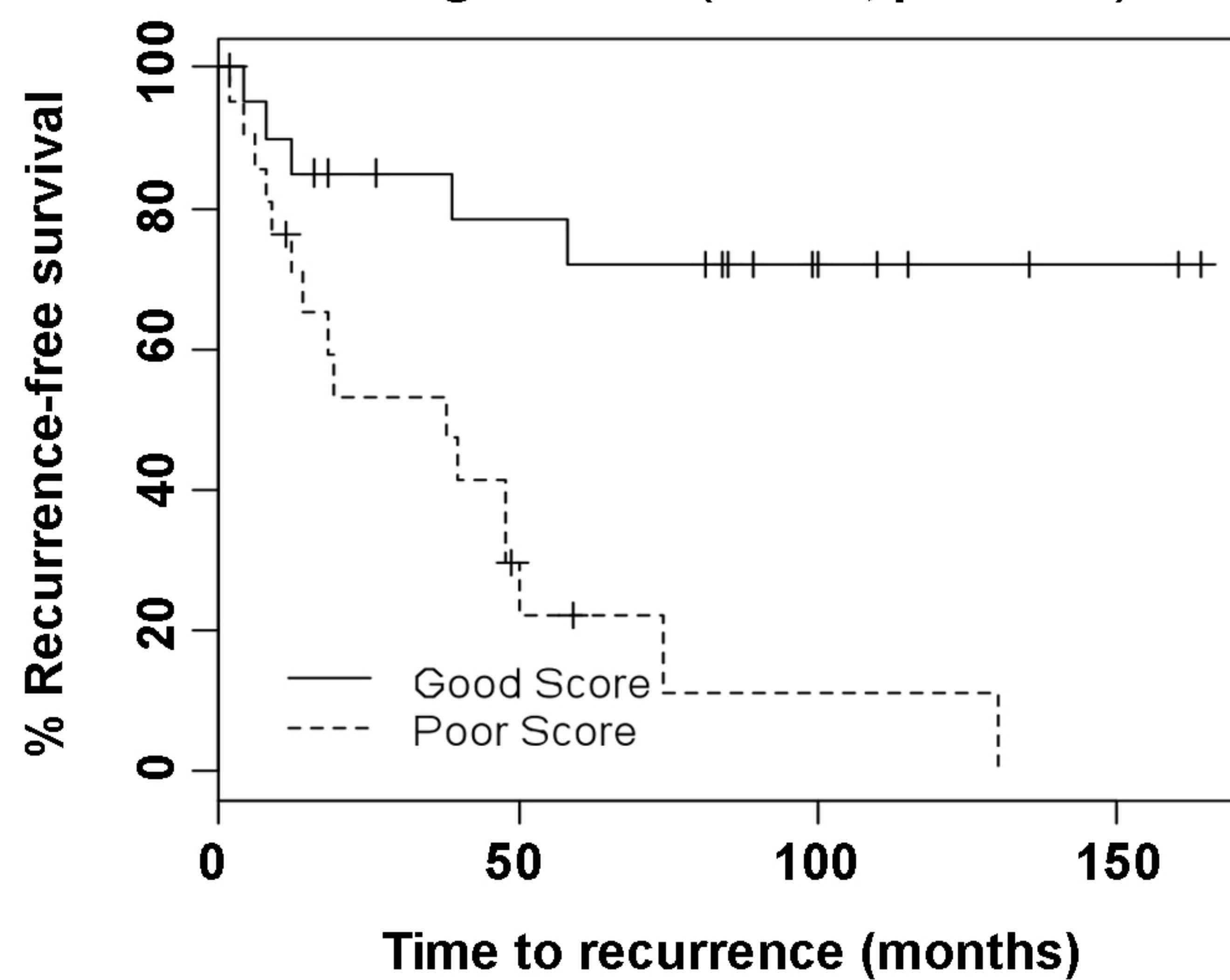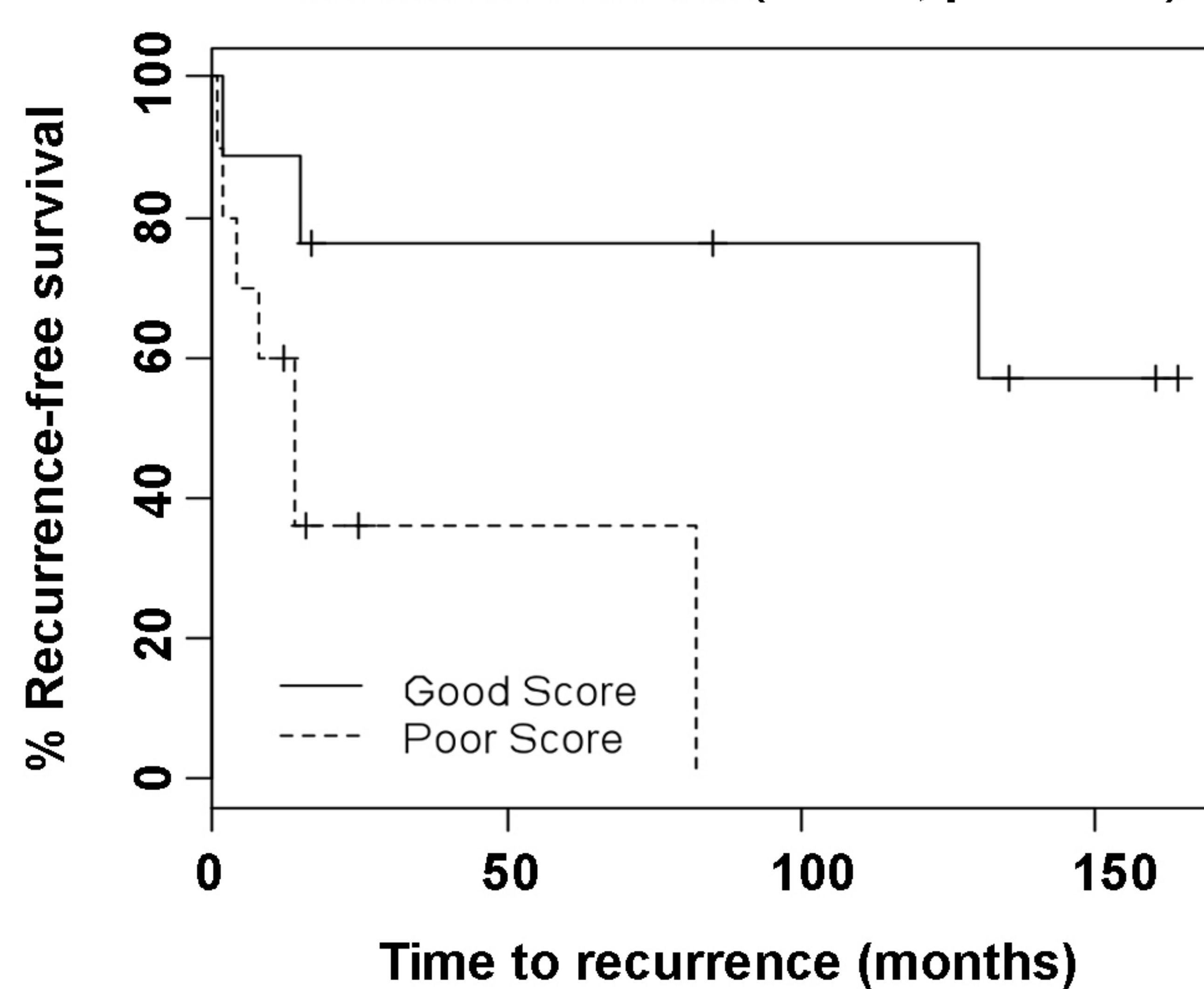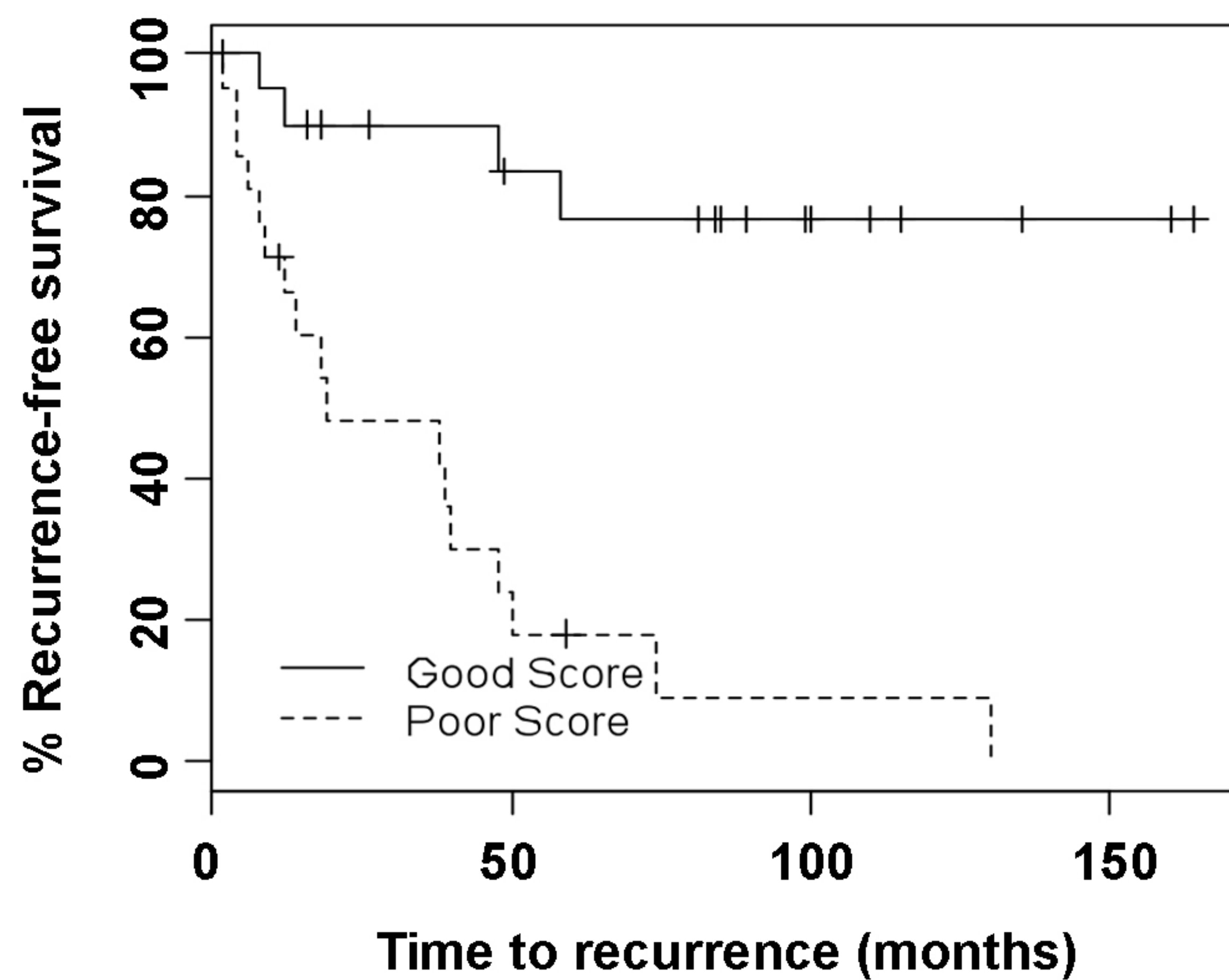| Training Set	Combined mRNA/miRNA panel	Clinical Information Only
All Cases (n=56)	**0.002**	0.262
Gleason score 7 (n=37)	**0.004**	0.136
Validation Set	Combined mRNA/miRNA panel	Clinical Information Only
All Cases (n=35)	0.078	0.193
Gleason score 7 (n=19)	**0.023**	0.080

Table 5: Summary of p-values (Logrank test) of prediction of clinical recurrence on training and validation sets for the entire dataset and the subset of Gleason score 7 cases using the combined and mRNA/miRNA panel, all of which are adjusted for T-stage, PSA, and Gleason score, or using clinical information only. Significant p-values are indicated in bold.

FIGURE LEGENDS

Figure 1: **Prediction of biochemical recurrence in all prostate cancer patients using two biomarker panels, adjusted for clinical information.** **(A)** Kaplan-Meier analysis of the training set patients that were separated based on the mRNA panel described in Table 2. **(B)** Kaplan-Meier analysis on the validation cases using the mRNA panel. **(C)** Kaplan-Meier analysis of the training set using the combined mRNA and miRNA panel described in Table 3. **(D)** Kaplan-Meier analysis of the validation set using the combined mRNA and miRNA panel.

Figure 2: Prediction of biochemical recurrence in prostate cancer patients with a Gleason score 7 using two biomarker panels, adjusted for clinical information. **(A)** Kaplan-Meier analysis of the training set of Gleason score 7 cases using the mRNA panel described in Table 2. **(B)** Kaplan-Meier analysis of the Gleason score 7 cases in the validation set using the mRNA panel. **(C)** Kaplan-Meier analysis of the Gleason score 7 cases in the training set using the combined mRNA and miRNA panel described in Table 3. **(D)** Kaplan-Meier analysis of the Gleason score 7 cases in the validation set using the combined mRNA and miRNA panel.

A

Training data set (n = 61, p < 0.001)

% Recurrence-free survival

Time to recurrence (months)

Good Score
Poor Score

B

Validation data set (n = 35, p = 0.010)

% Recurrence-free survival

Time to recurrence (months)

Good Score
Poor Score

C

Training data set (n = 61, p < 0.001)

% Recurrence-free survival

Time to recurrence (months)

Good Score
Poor Score

D

Validation data set (n = 35, p = 0.013)

% Recurrence-free survival

Time to recurrence (months)

Good Score
Poor Score

A Training data set (n = 42, p < 0.001)

B Validation data set (n = 19, p = 0.027)

C Training data set (n = 42, p < 0.001)

D Validation data set (n = 19, p = 0.010)

Good Score
Poor Score

% Recurrence-free survival

Time to recurrence (months)